

Distribution of Earnings Review - Report of the scoping exercise

- Summary
- Background and Introduction
- Conduct of the scoping exercise
- Content of surveys including earnings data
- Methodologies
- Outputs and dissemination strategy
- Processes
- Emerging structure of the quality review
- Annex A.

1. SUMMARY

As part of the National Statistics programme of quality reviews of key outputs, a review has been initiated into "the distribution of earnings, low pay and the New Earnings Survey". As a first step, an exercise to determine the scope of the review was conducted which, following the recommendations of the National Statistics Labour Market Theme Working Group for the conduct of quality reviews, looked at the content, methodologies and output of surveys which include earnings data. Following consultations with a range of users, eight areas for review emerged:

- i. User needs
- ii. Definitions and Concepts
- iii. Outputs
- iv. Methodology
- v. Revisions policy
- vi. Data linking
- vii. Processes
- viii. International

It is proposed that these eight headings form the basis for the work programme of the Quality Review. The next step is a Work Planning exercise which will further subdivide these eight areas in greater detail in order to prioritise them and translate them into identifiable chunks of work for allocation to review teams.

2. BACKGROUND AND INTRODUCTION

The White Paper, 'Building Trust in Statistics' (Cm 4412), called for

"a programme of thorough reviews of key outputs, at least every five years, with the involvement of methodologists and outside expertise, as appropriate;"

Following the establishment of National Statistics in June 2000, the responsibility for planning the programme of reviews has been taken up by the National Statistics Theme

Working Groups. The remit provided by the Labour Market Theme Working Group for its own programme of Quality Reviews states:

"Each review should address the following questions:

- Do the outputs meet user needs and are the outputs of adequate quality (the various attributes of quality for example: relevance, accuracy, timeliness, accessibility and coherence, should each be considered); can the information available to users on quality be improved?
- Is the methodology used appropriate and in line with best practice?
- Are the inputs required appropriate and fit for this set of purposes, and is the burden on data suppliers justified?"

At its meeting on 4 May 2000, the Labour Market Theme Working Group proposed a quality review entitled 'Earnings distribution, low pay and the New Earnings Survey'. This was intended to take forward a planned triennial review of the New Earnings Survey (NES), but widen it out to consider the context of earnings statistics more generally. The review would consider the NES alongside earnings data from other sources, particularly the Labour Force Survey (LFS); smaller surveys such as the Family Expenditure Survey would also fall within the scope of the review, although it should concentrate mainly on the earnings aspects of the major surveys. The Average Earnings Index, and the Monthly Wages and Salaries Survey on which it is based, are subject to their own detailed programme of reviews and hence fall outside the scope of the Earnings Distribution review except for the context they provide.

The first stage in the conduct of the review is a thorough scoping exercise which examines current user needs for earnings data, and asks how far the available surveys meet those needs in terms of the questions asked, the methodologies used and the outputs produced. This document reports on the results of that exercise.

3. CONDUCT OF THE SCOPING EXERCISE

The central data gathering exercise on which this report is based is a series of interviews with ONS staff and representatives of outside organisations concerning the conduct of ONS surveys which include earnings data and the needs of users.

Customers were represented by a range of organisations including government departments and the commercial and business sector. A full list of those who took part in the consultation is given in Annex A. Due to time restraints views were sought from representative bodies, such as the CBI rather than individual companies. Representatives of some sectors, such as local authorities, were unavailable during the course of the scoping exercise but the intention remains to seek their views during the course of the review itself.

Following the remit provided by the Labour Market Theme Group, the consultations centred around

- The content of surveys: questions asked in surveys such as the New Earnings Survey and LFS, and the concepts and definitions used
- the methodologies used
- the dissemination strategy, including the presentation of outputs.

One other aspect of the review is the internal processes that ONS uses in the transformation of data, from sending out questionnaires to producing final reports. These processes are in the main invisible to customers and formed part of a separate consultation of ONS staff.

4. CONTENT OF SURVEYS INCLUDING EARNINGS DATA

4.1 Concepts and definitions

Concepts used in the NES were queried, with some respondents commenting that they can seem out of date. Those concepts mentioned were:

- the concept of adult rates of pay as opposed to training rates. Nevertheless, training rates of pay are relevant to National Minimum Wage legislation and the problem may be one of nomenclature in that there is no longer a link between training and age;
- the definition of full and part time in terms of hours worked, especially given the emergence of more flexible working patterns;
- the manual / non-manual distinction was blurred in some occupations. Some occupations classed as non-manual could nowadays involve more physical labour than many manual jobs. The emphasis should be placed more on skill levels.

One respondent felt that a harmonised definition of 'employee' should be used across National Statistics. Currently, different definitions appear to be used in the NES, LFS and Workforce Jobs series. There is a more general need to ensure that concepts and definitions are as consistent as possible between surveys, where no overriding reasons for differences can be identified.

4.2 Content

4.2.1 Business versus household surveys

A number of issues were raised that arise from the design of surveys such as the LFS and NES, and the relative advantages of business versus household surveys. The possible range of variables, the perceived accuracy and the achievable sample size are some of the factors that differ between the two types of survey, giving rise to advantages and disadvantages when comparing one with the other.

The NES tends to be preferred for its large sample size, perceived accuracy and reliability, and for the possibility of constructing long time series of data. The LFS is used for its wider range of variables. Although the LFS provides more information about individuals, its sample size prohibits detailed breakdowns such as are needed to

analyse the qualifications of the low paid or the effect of National Minimum Wage (NMW) legislation on different groups, for example.

Ideally, respondents would like the best of both worlds - a large sample size and a wide range of variables. However, there was an appreciation of the difficulties in collecting personal information on an employer-based survey.

Some of the suggested additions to the NES were:

- Ethnic origin.
This would be difficult to collect from employers
- Qualifications.
Again, difficult for employers to determine
- Skill levels.
How this could best be measured has not yet been determined
- Collective bargaining.
The existing information on collective agreements might be enhanced by incorporating some questions from the Workplace Industrial Relations Survey, providing a wider range of data, with increased accuracy and on a more regular basis.
- Hourly rates of pay.
The addition of an hourly rate question, similar to that used in the LFS, would enhance NES data.

Naturally, any consideration of larger sample sizes or an increased range of variables must also take into account the costs involved, both to the public purse in running the surveys and the costs to businesses taking part.

4.2.2 Components of pay

Components of pay and additions to basic pay were areas of interest for many respondents. The current categories in the NES were not seen as wide enough to capture the range of modern remuneration practices. In particular, there was a fear that some categories were used as a 'catch-all' for a diversity of pay practices which didn't fit elsewhere. A number of respondents felt that additions to basic pay on the NES should be measured on a consistent basis with the AEI. Some guidance from ONS on the relative merits of the AEI and NES as sources of information on additions to basic pay would be welcomed.

Examples of current remuneration practises that respondents would like to see identified included:

- benefits in kind
- pensions
- share options
- bonuses based on individual merit to be distinguished from company wide payments
- tips
- accommodation
- redundancy payments

There was considerable interest in the Structure of Earnings Survey, to be conducted in 2002, which may act as a pilot for new questions on components of pay.

The Confederation of British Industry suggested that payment systems in general would be a fitting subject for a survey. It was probably more the kind of survey that CBI or the Chartered Institute of Personnel and Development might conduct; nevertheless, the possibility of ONS co-operation would be met with interest. Ideally, such a survey could be completed in time to feed its findings into the planning of the SES.

4.2.3 Pay settlements

The Trades Union Congress were particularly interested in pay settlements, and suggested that ONS might consider a survey to run in parallel with the NES and AEI. Other respondents drew attention to commercially available settlement data. The review could consider the extent to which data on pay settlements already exists and the possibility of comparing this with earnings growth measures from the AEI and NES.

4.2.4 Industry and occupational coding

There are differences between the industry coding on the NES and that on the LFS which are known to arise from the different ways in which the two sets of data are collected. Related to this is the question of whether there are similar differences between surveys in earnings by occupation. The review should investigate the extent to which this occurs and the possible causes.

5. METHODOLOGIES

5.1 Correction for bias

It was widely agreed that the NES should be grossed or weighted to correct for undersampling, although it was recognised that this could not compensate entirely for the lack of NES coverage of employees with earnings below the PAYE threshold. In fact, current ONS work programmes include two areas, imputation and weighting for non-response and the development of a methodology (using NES and LFS data together) for estimating the numbers of employees earning below the National Minimum Wage. These developments should help address a number of concerns held by respondents, in particular the undersampling of small firms in the NES. The review will report on progress in both these areas and consider what further work is needed, including an investigation of whether the NES sample can be extended in order to cover more comprehensively employees with earnings below the PAYE threshold.

5.2 Accuracy

The development of a methodology for estimating the numbers of employees earning below the National Minimum Wage had raised questions about the quality of LFS estimates, which were generally not considered robust enough for work on the NMW. The introduction of the LLFS was welcomed as a route to improving the accuracy of estimates through the sample boost.

The quality of data on hours worked was considered a suitable topic for the review to cover, since hourly earnings is reliant on accurate measures of hours worked. Other

aspects of estimating the numbers earnings below the NMW - the central estimate methodology, the treatment of proxy responses in the LFS - would also be eligible for inclusion in the review.

Although the accuracy of the LFS was questioned, it was the only source of earnings data for some groups, such as ethnic minorities. One respondent suggested that earnings data in the LFS should be asked of every wave in every quarter, thus increasing the sample size available for analysis.

Regarding the NES, one respondent suggested that the need for such a large sample should be addressed by the review. However, most users were in agreement that the sample size of the NES was necessary to allow the detailed occupational and sectoral breakdowns that they needed.

5.3 Linked datasets

A number of respondents would like to see some linking of different datasets. The use of national insurance numbers presents a wealth of opportunities. Datasets which could be linked to the NES that were mentioned were

- Claimant count unemployed
- Qualifications and training data
- DSS benefit data
- Home Office prison records

Many respondents asked why previous work linking the NES to JUVOS data through National Insurance numbers had not become part of the regular outputs. It is understood that this work is still continuing, and the review should therefore look at progress to date. Possible uses for such a linked series include examining the effects of unemployment spells on earnings levels, or conversely determining whether the low paid experienced more frequent periods of unemployed. The latter could help determine whether a bias might be introduced if the low paid were more likely to disappear from the sample between February and April.

5.4 Reconciliation

There was a call for reconciliation to be given a high priority. Reconciliation of the LFS and NES in particular would help address the problems of undercoverage of low income groups encountered from the PAYE sampling frame. Users felt that they did not have access to the data in sufficient detail to carry out their own reconciliation exercises and therefore looked to ONS to take the lead. ONS is already conducting reconciliation work on earnings; the review should look at existing programmes and identify areas where further work is needed.

5.5 Coding

Discontinuities, particularly in the coding of industry, occupation and agreement and in the recording of components of pay, were considered a major barrier to the construction of long time-series. A number of respondents called for the construction of a back-coded dataset.

Resolution of problems surrounding geographical codings were a high priority with those who would like to access NES data through NOMIS. The possibility of accessing data by the postcode of an employee's residence was also greeted with interest.

5.6 Timing of surveys

The question of the timing of the NES, and the effects that bonuses would have on earnings levels, was raised. Delayed bonuses were seen as particularly problematic, and it was felt that the review could look at ways to quantify and correct for the effects on earnings that might arise.

5.7 International comparisons

Although it was recognised that comparing data sources from different countries was fraught with difficulties, there was considerable interest in international comparisons. Comparisons of the distribution of earnings would be of interest even if they were fairly minimal - reporting the top and bottom decile, for example. Interest was expressed in examining the effects of minimum wage legislation in other countries, components of pay (whether bonus systems are extant in other countries) and differential effects of collective bargaining. While the conduct of comparative studies would be out of scope, the review could address the feasibility of such studies by investigating the extent to which UK data sources are compatible with other countries and the interface with Eurostat.

6. OUTPUTS AND DISSEMINATION STRATEGY

6.1 Electronic dissemination

Government users wanted access to microdata so that they could conduct their own analyses, construct time series and take full advantage of the flexibility that this would give them. They wanted the data in electronic format and ideally at no cost. Current policy on access and charging appears to be confusing. While the review could highlight some of the problems, policy is likely to be determined by the National Statistics Code of Conduct and falls outside its scope.

Electronic dissemination was important to most users, but the preferred format varied. Some wanted access via the web site. Government departments and the Bank of England would welcome the publication of long time series data via the Central Shared Database (CSDB). Users who had access to the LFS via NOMIS or SPSS-MR would welcome similar access to the NES.

Concerning existing sources, StatBase was difficult to use for some respondents as there is no cross-linking between headings. A user might search for earnings data under the 'labour market' or 'economic' themes, for example, and having chosen the wrong route would find it difficult to get on the right track. The search facility produced too many or too few hits. Cross linking, so that earnings data could be found under economic and labour market headings, would be the easiest solution for users.

A recent LFS historical series on CD ROM was cited as an example of good practice which the NES could usefully emulate. Other users would like panel data, or a linked

dataset, that would allow them to construct their own time series. It was also felt that earnings data from the AEI, NES and LFS could be gathered together in one (electronic) source. A CD-ROM version of the published tables was seen as potentially useful, but rather less so than panel data or linked data.

6.2 Printed output

Printed output from the NES was not considered to be user friendly, although it contained a wealth of detail for those familiar with the six volumes of the published data. A common complaint was that the tables all looked the same, so that the user would find him/herself on the wrong page. A change of font for table headings was suggested although the layout of the tables is considered to be better than it was in the past. The levels of aggregation were not always what the user wanted; for example, the highest level of aggregation might be full-time males when the user wanted all males in employment.

One theme that emerged was the utility of commentary from ONS. A publication along the lines of Social Trends ('Earnings Trends', perhaps), with analysis and graphs, would be a welcome addition to the printed output. The independence of National Statistics should be exploited; users felt that ONS should publish analyses to ensure that they were accessible and not written by someone with an 'axe to grind'. Existing commentary and analysis in Labour Market Trends was considered useful, especially when regular analyses were conducted such as those using the LFS - ethnic origin and the labour market was given as one example, where the regular appearance of an article could be relied upon. Similar analyses of earnings data would enhance the existing published outputs, along with an 'Earnings Trends' publication. Such analyses would help some users by doing some of their work for them, but the main expected benefit would be that it would illustrate what it was possible to do with earnings data. It would also give earnings data a higher profile in the research community.

The First Release briefings were considered to be very good and most users found them fit for purpose. They contained the right information and their availability in PDF format from the website made them accessible and timely. Those who had attended press briefings were similarly impressed.

The frequency and timeliness of the NES was mentioned, with many users asking for increased frequency. Methodologies were discussed, including the use of smaller interim surveys and the use of the LFS to interpolate monthly movements in earnings.

6.3 National Accounts

Earnings data will be used for National Accounts and the review could usefully address the question of whether the data can be produced in the formats which will be required. The Economy Theme Working Group is responsible for quality reviews of National Accounts, and the Distribution of Earnings Review should keep track of developments in this area.

7. PROCESSES

The processing of the NES for the year 2000 had been subject to new procedures consisting of the introduction of common software for data input and validation. There had been problems with the capacity of the server which meant that throughput was slower than in previous years, which had threatened the publication timetable. It is understood that these problems will be rectified in time for the 2001 survey, but nevertheless there will be lessons to learn and the review could usefully summarise these.

Processing the data for each survey is a complex procedure involving a number of databases in addition to the common software. The issues of concern to users are timeliness and accuracy. To address these concerns, the review could consider a number of questions, including

- the validation checks in use - are they necessary and effective?
- the procedures for checking data - are they efficient and accurate?
- the reasons for 'late returns' - is there scope for reducing the number of responses which are not included each year?
- the procedures for following up non-response - can improvements be made to the numbers of non-responses identified and included?
- the scope for extending 'special arrangements' - can improvements in efficiency be achieved by increasing the number of respondents supplying direct returns?

8. EMERGING STRUCTURE OF THE QUALITY REVIEW

Many of the users interviewed had strong and consistent ideas on a number of topics suitable for inclusion in the Review programme. They can be grouped under eight headings.

i. User needs

Current needs for the NES and LFS - sample size and accuracy versus detail
Content of the surveys
Uses and fitness for purpose
Related data needs, eg labour costs, pay settlements, collective bargaining, hours worked

ii. Definitions and Concepts

Components of earnings and links with the Structure of Earnings Survey
Definitions of employee
Relevance of current distinctions - adult rates of pay, full / part time, manual / non manual
Payment systems and benefits in kind
Hours of work basis used - usual and actual hours
Industry classification on IDBR - local unit versus reporting unit
Occupational differences in earnings arising from differences in classification

iii. Outputs

Electronic dissemination; current progress in dissemination via NOMIS, the NS website (including Statbase), the CSDB and SPSS-MR

Printed output - how to make it more user friendly

Commentary, analysis and added value - 'Earnings Trends'

Single access point for earnings data from NES, LFS and AEI

Timeliness

Historical supplement

iv. Methodology

Imputation

Grossing

Reconciliation

Estimates of persons earning below the National Minimum Wage

Sample design

Accuracy; sample sizes, treatment of proxy responses in LFS

v. Revisions policy

Frequency of surveys

Discontinuities, construction of a back-coded time series

Use of AEI or LFS to interpolate revisions to earnings based on NES between surveys

Dealing with late returns

vi. Data linking

Other sources of data

Panel data

Use of linked data to investigate bias

vii. Processes

Data validation

Response chasing

vii. International

Comparison of earnings distributions

Components of pay in different countries

Distribution of pay

Collective bargaining

In conclusion, it is proposed that these eight headings form the basis for the work programme of the Quality Review. A Work Planning exercise will translate the eight areas into identifiable chunks of work for allocation to review teams. It is not envisaged that the work areas should have equal weights; the planning will include prioritisation of work and consultation with review team leaders to determine available resources. This will result in a draft of the timetable and resources required for each section.

ANNEX A

Organisations and individuals taking part in the consultation

1. Government Departments

Office for National Statistics	Barry Werner Claire Nicholl Dave Austin Pete Douglas
Department for Education and Employment	Dennis Allnutt Iain Bell
Department of Trade and Industry Low Pay Commission	Janet Dougherty Joanna Selden Ewa Kmietowicz
Her Majesty's Treasury	Chris Kelly Nicholas Vaughan Duncan Melville Tim Riddington

2. Other organisations

Bank of England	Mike Joyce Fergal Shortall
Confederation of British Industry Incomes Data Services	Peter Haslett Alastair Hatchett Nicola Allison Sally Brett Ken Mulkearn
Trades Union Congress	Iain Murray Ian Brinkley
National Institute for Economic and Social Research	Gary Young

3. Organisations unable to take part in the consultation at this stage

Local Government Association London Chamber of Commerce and Industry	Juliet Whitworth Paul Valentine
---	------------------------------------

If you wish to send us your comments on any document on this page please email the National Statistics Quality Review Team and your comments will be forwarded to the relevant person as appropriate.
