Re-designing the UK's Annual Structural Earnings Survey

Mark Pont

Office for National Statistics
Cardiff Road, NEWPORT, UK, NP10 8XG. Email: mark.pont@ons.gov.uk

Abstract

The methods underpinning the UK's annual structural earnings survey – the New Earnings Survey (NES) – have remained largely unchanged since the survey's inception in 1970. A recent major review of earnings statistics in the UK, along with gradual changes in the labour market in recent years have led to a re-design of the NES.

This paper describes some of the shortcomings of the NES and then presents the methodological work undertaken to re-design the survey. The methodological work covered several topics:

- better use of administrative data;
- development of the sampling design (which exploits a dual frame approach to reduce under-coverage of the survey);
- improved data collection procedures, including an intensive follow-up pilot;
- better procedures to analyse the characteristics of people who change jobs; and
- a weighting strategy to account for both non-selection and non-response.

Experiences from pilot work are described along with plans for full implementation in 2004.

Key words: distribution of earnings, dual frame survey, sampling design, weighting

Introduction

The New Earnings Survey (NES) is the main source of data on the distribution of earnings in the UK. The NES has been run, broadly in the same form, in every year since 1970. It measures the earnings of employees in employment in the UK across the whole economy in April of each year. The NES is conducted by the Office for National Statistics (ONS) for Great Britain and the Department of Enterprise, Trade and Investment in Northern Ireland (DETINI).

Most of the 1 per cent random sample of employees is drawn by the Inland Revenue (the UK's government department responsible for collecting income tax) from the Pay As You Earn (PAYE) system by reference to the last two digits of an employee's National Insurance number. A small proportion of the sample is identified, using the same selection criterion, directly by employers who return their information electronically to ONS. The sample design creates what is effectively a panel of employees. The sample size in recent years has been about 240,000 employees.

Employers are required (under the Statistics of Trade Act 1947 and the Statistics of Trade and Employment (NI) Order 1988) to return data for each of their employees who fall within the NES sample, showing details of, among other things, their earnings, hours worked and a description of their occupation.

The survey's outputs are used:
- for national minimum wage evaluation, local authority resource allocation, pay review purposes, and labour market policy assessment (annual structural data);
- to meet European Union Regulations (biennial structural data on earnings and labour costs); and
- for policy evaluation, modelling and impact assessment (longitudinal data).

Shortcomings of the NES

When the NES was first run, the outputs were designed to meet the need for data on the average weekly wage of full-time employees, and in that respect the use of the PAYE system as the sampling frame was appropriate. However, users now need to look at employees across the whole economy, which includes part-time workers and workers on low rates of pay and with low weekly wages. In particular, since the introduction of a National Minimum Wage, precise estimates are needed of the

number of people earning low hourly rates. Thus, the continued use of the PAYE system alone means the NES is unable to satisfy users' requirements as adequately as it once did. The exclusion of so many employees (over a million employees may be missed from the population – the Labour Force Survey estimates approximately 25.6 million employee jobs) means the output no longer adequately yields a measure with a relevant concept. Therefore, although the NES can produce precise estimates of pay in terms of variance, bias means the statistics could be misleading. Bias is also likely to vary from industry to industry as the excluded employees are generally concentrated in a few industries.

Further shortcomings that could affect the quality of estimates from the survey include the questionnaire which had not been subject to a thorough review but had been built up incrementally over time, the response chasing strategy which appeared to exacerbate non-response bias, and the lack of weighting in estimation of results to compensate for non-response bias.

Addressing the shortcomings

The shortcomings were identified by a major review of earnings statistics – The Distribution of Earnings Review (DoER, ONS 2002). In order to address the shortcomings, a major project was established in ONS towards the end of 2002 to review the methods and operations in the NES. As the survey will be substantially different from 2004 a new name has been proposed, but not yet adopted, and for sake of clarity we continue to refer to the redesigned survey here as the NES. The main aims of the project are to:

improve coverage of the survey by collecting data relating to employees who are not part of a PAYE scheme and those who have recently changed employer;
review the questionnaire;
evaluate the impact of more intensive follow-up of non-responders;
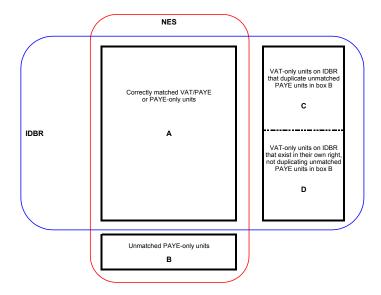produce weighted estimates and consider stratification of the sample; and
calculate standard errors of some of the main output variables, including quantiles and changes in quantiles.

Each of these topics is now addressed in more detail.

Improving Coverage

Compared with the standard business survey population used by ONS – the Inter-Departmental Business Register or IDBR (Perry 1995, ONS 2001) – the NES population is deficient. The IDBR is based on two main sources – PAYE employers and traders registered for Value Added Tax (VAT) – which are matched together. Therefore, any employees working for VAT-registered traders will be automatically excluded from the NES. In addition, the matching process is not perfect, which further complicates the population definition.

The following diagram illustrates how the NES frame population from PAYE and the IDBR population (from the combination of VAT and PAYE registrations) overlap. Category A contains those units that are correctly structured on the IDBR and the employer is part of the NES population. Category B represents a small number of employers where the employee is within the NES population, but the PAYE unit itself has been stored off the main IDBR because of a failure to match it against the relevant VAT unit. The characteristics of these PAYE units are such that we expect a match against the relevant VAT unit, but because of imperfections in the name and address of each unit we have failed to find the match. In such cases, the VAT unit is given priority because, in general, the VAT units have better quality information. In order to minimise duplication these records do not form part of the main sampling frame. Categories C and D together represent businesses on the IDBR that are based only on VAT records. In most cases these are businesses that are not large enough to have an employee (category D) but there are some cases where the record represents a failure to match against a PAYE unit (category C); categories C and D are indistinguishable on the business register. Conceptually at least, the levels of coverage of B and C are equivalent.

NES

IDBR

Correctly matched VAT/PAYE
or PAYE-only units

A

VAT-only units on IDBR
that duplicate unmatched
PAYE units in box B

C

VAT-only units on IDBR
that exist in their own right,
not duplicating unmatched
PAYE units in box B

D

Unmatched PAYE-only units

B

In addition to the under-coverage represented by category D, which are generally termed *VAT-only employees*, two other major sources of under-coverage have been identified. Employees who work for a business that runs a PAYE system, but the employees themselves are not included in the system because they don't earn enough to pay tax ("off-PAYE employees"); and those who have changed employer, and about whom it has previously been difficult to collect information ("EC3s").

In addition, some businesses are registered neither for VAT nor PAYE. These are outside the scope of the IDBR, and consequently the NES. It is possible that some employees will be missed from these businesses, and that these will remain a potential source of bias.

VAT-only employees

In order to assess the extent to which VAT-only businesses had employees and the ease with which we could collect data about them, a pilot survey of 2000 such enterprises was conducted during 2003. Prior to the pilot, a small pre-pilot was undertaken during which we telephoned about 100 such businesses. The results of the pre-pilot were generally inconclusive, but gave enough indication that it would be worthwhile to proceed with the main pilot.

The pilot survey was despatched during April 2003 (the same time as the main sample was despatched). A similar questionnaire to the main survey was used with the addition of an extra sheet that guided the employer to answer questions relating to whether it operated a PAYE system, or had employees that were not part of a PAYE system.

At the time of writing, the data from the pilot survey are still being validated and cleaned, but the pilot suggests that such enterprises could represent up to 200,000 employees. These are split between employees who belong to PAYE schemes that haven't been correctly linked (these are represented by the main NES sample in area B) and those that don't belong to a PAYE scheme. Once the data have been fully cleaned, further analysis will be undertaken to inform the sample design for 2004.

The pilot study showed that it is viable to collect data from such businesses, although some refinements to the data collection and validation processes are required to improve the quality of the data.

Off-PAYE employees

These employees are more troublesome because head offices will typically be unaware of the precise details (hours, pay etc) of such employees, whose 'contracts' are generally arranged locally. Such employees will typically be in the service sector, working for example as weekend bar staff in restaurants, bars and hotels. In order to collect data about these employees, it is necessary to contact the individual sites (local units) operated by each employer. A small telephone pilot of hotels and restaurants is presently underway, and is showing the existence of some such employees, but also highlighting difficulties in collecting data about them. The work is ongoing.

EC3s

For around 10 per cent of responses to the NES the employer tells us that the particular employee no longer works for them, but is likely to be working elsewhere. Clearly, because such employees are no longer with the employer at the time of the survey, it is impossible for that employer to respond with any data. This is expected to introduce bias into results, as employees changing employer are likely to exhibit different patterns of both levels and change in pay from employees who remain with the same employer. We know that many of the cases are due to lags in the administrative processes that drive the production of the sample. It is worth noting that some of the responses we get from businesses explain that the employee actually left the employing business many years previously indicating an element of 'deadwood' in the sampling frame, representing an element of overcoverage in the frame.

In order to overcome these difficulties a pilot is being run during 2003 whereby 3000 or so EC3 responses will be matched against a later version of the sampling file to determine whether the employees are now recorded against a different employer from that previously shown. We intend to send a new questionnaire to each of the new employers to collect the relevant information about each employee. We expect this approach to be only marginally successful since it relies on the new employer having been recorded in the administrative system, and the date of data collection is later than for the main survey which could cause delays in production of results. An additional, as yet untested, proposal is that we could send some blank forms with the main despatch (the main forms are pre-printed with the appropriate employee's name and NI number) and ask employers to complete one for each eligible employee for whom we haven't sent a pre-printed form, but this needs further evaluation as it adds significantly to operational costs.

Questionnaire Redesign

DoER recommended that the NES questionnaire be redesigned for 2004. The redesign is being taken forward through a specific research and development programme that forms part of ONS's greater attention to the rigorous design processes that are required for business survey questionnaires. The scope of the programme is to: obtain a specification of the data requirements of NES customers; identify through an 'expert review' process, questions and aspects of the NES questionnaire which require redesign; and redesign and test the questions and the questionnaire, including questions that will not be included in the survey every year.

To date users have been asked for an up-to-date assessment of their requirements for data from the survey and an expert review of the questionnaire has been carried out. Following this, a revised questionnaire has been written and is currently being tested in the field.

Intensive Follow-Up of Non-Responders

Previous work showed some differences between the response values for late responders and those relating to employers that responded punctually. It is likewise expected that non-responders may be different from punctual responders.

In order to estimate the effect of this, an intensive follow-up (IFU) exercise is planned to take place after the regular data collection has ended in respect of the 2003 survey. A sample of non-responding employers will be chosen at the appropriate time and these will be telephoned several times in the hope of improving the response rate for this subset closer to 100 per cent. Depending whether the employees of these businesses have different characteristics, we will decide the strategy to deal with this in future.

Sampling and Weighting

We considered whether the sample should be stratified. We investigated how this could be done given that the variables most useful for stratification were unavailable on the sampling file. A hybrid system was proposed whereby the occupation and region information would be taken from previous responses to the NES where known. Given the panel nature of the NES sample and the large overlap from year to year, this would allow a large proportion of the sample to be allocated to the correct stratum. Given that the constraint of a maximum sampling fraction of 1 per cent would remain, the purpose of the stratification would be simply to identify areas of the population with low variability where sampling could be lighter in order to reallocate the survey's resources. However, it is now expected that such a design will be introduced in 2005, with the sampling in 2004 proceeding as in previous years.

One of the major deficiencies of the NES has been the lack of weighting in the production of results. Therefore, classes of employee with low response rates will be under-represented in estimates, leading to possible biases in the survey. Our analysis showed wide variations in the response rate between different occupational groups defined by ONS's Standard Occupational Classification 2000 (TSO 2000).

In investigating which benchmark totals to use for calibration, we found that the variables most closely correlated with earnings in the UK are occupation, age, sex and geographical location (salaries are much higher in and around London than the rest of the country, which led to two geographical strata). We therefore wished to calibrate against known population totals broken down by as many of those variables as possible. The Labour Force Survey (LFS) produces estimates of the number of people (including the number with second jobs) for each of these categories. The number of additional jobs is estimated to be about 100,000; a value small enough not to be of concern in the context of calibration.

We have therefore adopted an estimation system where the design weight reflects each individual employee's probability of selection and calibration weights are calculated against benchmark totals for each of the 500 calibration groups (25 sub-major groups of the ONS's Standard Occupational Classification 2000 × five age groups × two sexes × two geographical regions). The weight to be applied to each employee is a combination of a design weight and a calibration factor.

The design weight is defined as $a_i = \dfrac{N}{n'}$ where n' is the number of responses to the survey. Since N is unknown, we estimate this as $\hat{N} = 100n$, where n is the number of employees on the initial sample file. The calibration factor is then defined as $g_i = \dfrac{N_{h(ps)}}{\sum\limits_{j \in h(ps)} a_j}$ where $N_{h(ps)}$ is the calibration benchmark total from LFS in calibration group (or post-stratum) $h(ps)$.

This formulation holds under stratification or when including further employees arising from, for example, the VAT-only pilot as long as a design weight can be calculated for these employees.

Weights have been calculated for NES 2002 data using Statistics Canada's *Generalized Estimation System* (GES) (Estevao et al 1995). These have been used to successfully create weighted estimates for a range of variables.

One area of concern we had to address was that the calibration totals are themselves estimates from a sample, so NES estimates may be adversely affected. Two variants of LFS data set are available – quarterly and annual. The annual data sets are formed from an amalgamation of the quarterly data sets, with the addition of additional information from various top-up samples that are used to improve the quality of estimates for small areas. Because of the way the data sets are put together, the annual data sets differ from the sum of the relevant four quarters.

The quarterly data sets are available within six weeks of the end of the quarter in question. The annual data sets are more desirable from the point of view that the estimates therein are more accurate, but they are less timely. Furthermore, it could be argued that choosing the quarterly data set closest to the survey date might be more appropriate. Our work showed very little sensitivity to the particular data set being used; the range of estimates of average wage generated from the various candidate data sets being less than 1 per cent of the value itself. Timely availability of the annual data sets might necessitate use of the quarterly data; further work is ongoing to explore the possibilities.

Calculation of standard errors

Prior to the redesign project, some standard errors had been produced for estimates of unweighted means. However, no reliable standard errors were available for estimates of changes in the mean, for the levels or change of any of the quantiles or for weighted estimates. A major part of the project was to be able to calculate such estimates in order that users can be informed of the quality of NES outputs.

The development of estimates of standard error for means and changes in means is relatively straightforward under a weighting scheme and have been calculated using similar methods to those developed for other ONS business surveys. More difficult is the issue of standard errors for medians (and other quantiles) and yearly changes in such estimates. To estimate the

standard error of the median we adopted the approach in Särndal et al 1992, p 198 which approximates the distribution around the estimate of the quantile in question.

The cumulative density function, $\hat{F}(y)$, is estimated by cumulative addition of the relative weights. A particular percentile, x_p, is estimated by $\hat{x}_p = \hat{F}^{-1}(p)$. A confidence interval is then determined based on the assumption (which is justified for large samples) that $\hat{F}(x_p)$ is approximately normally distributed around its expected value, $F(x_p) = p$.

$$C_1 : p - t_\alpha \{V[\hat{F}(x_p)]\}^{1/2} \qquad C_2 : p + t_\alpha \{V[\hat{F}(x_p)]\}^{1/2}$$

where t_α is the standard normal deviate for a two-sided confidence interval of size $1 - 2\alpha$.

For the weighted data set, $V[\hat{F}(x_p)] \approx \sum_h \dfrac{p(1-p)W_h^2(1-f_h)}{n_h}$, where h refers to stratum h, n_h is the number of observations in stratum h, N_h is the stratum population, f_h is the sampling fraction and $W_h^2 = \dfrac{N_h}{N}$, where N is the population total.

The standard error of a quantile is estimated from the confidence interval as $\dfrac{\hat{F}^{-1}(c_2) - \hat{F}^{-1}(c_1)}{2t_\alpha}$

This method produced credible estimates of standard errors, that of the median being smaller than the standard error of the mean. Estimating the standard error of the change in quantiles was more difficult, and development of the formulae wasn't possible in the time allowed for this project. As an alternative, bootstrap estimates of standard errors of changes in quantiles have been calculated and appear to be of acceptable quality.

One curiosity we encountered when estimating standard errors of quantiles was that some estimates of standard error would be markedly lower than expected. This arises because of 'clumping' at certain points of the income distribution corresponding, for example with annual salaries being whole numbers of thousands of pounds. This occurs both because that is the actual level of salary and also because of some approximation by the reporting business and occurs at 'round' values of annual salaries, weekly wages and hourly rates. This presents a dilemma for publication – should we publish the estimate of standard error for the quantile knowing that it is atypical, or do we publish some kind of average value that doesn't necessarily represent the extra accuracy in the particular period? Furthermore, in deciding the accuracy requirements to be used as the basic criteria to determine whether an estimate is accurate enough to publish at all, we need to consider whether a particular year's value is one of these atypically accurate values.

References

Estevao V, Hidiroglou MA and Särndal C-E (1995) Methodological principles for a generalized estimation system at Statistics Canada. *J. Off. Stat.* 11 181-204.
ONS (2001) *Review of the Inter-Departmental Business Register*. London: ONS.
ONS (2002) *Review of Distribution of Earnings Statistics*. London: ONS.
Perry J (1995) The Inter-Departmental Business Register. *Economic Trends* 505, November 1995.
Särndal C-E, Swensson B and Wretman J (1992) *Model Assisted Survey Sampling*. New York: Springer-Verlag
TSO (2000) *Standard occupational classification Vol. 1 Structure and descriptions of unit groups*. London: The Stationery Office.