

## FEATURE

Robert Hayes, Catrin Ormerod and  
Felix Ritchie  
Office for National Statistics

# Earnings: summary of sources and developments

## SUMMARY

UK earnings data are collected through many different channels and summarised using different methods. Hence, the stories told by aggregate figures may appear to conflict with one another. While the overall picture of earnings is similar, there are important differences in the detail.

This article draws together the various sources of earnings aggregates to investigate similarities and inconsistencies and to improve understanding of this data. In particular, this work highlights the benefit from bringing sources together at the unit level to enhance analytical capability and throw more light on the coherence between data sources.

In the last eighteen months, the Office for National Statistics has linked datasets using a range of innovative techniques to enhance analytical capability and carry out micro-level consistency checks. The richness of the linked datasets has brought out a number of new important results which are summarised in this article.

The Office for National Statistics (ONS) is the main source of statistical information on earnings in the UK, gathered by surveys of individuals or their employer. Some surveys which focus on topics other than earnings also include earnings information, mainly to allow descriptive analysis. Other government departments routinely collect information on earnings; for example, HM Revenue and Customs (HMRC) collects earnings data for taxation and income forecasting purposes. 'Earnings data: A brief guide to sources and outputs' describes these; a more detailed version of the article, 'Earnings data: Sources and outputs', is available on the National Statistics website.

The Earnings Analysis branch of ONS was set up in March 2005 to enhance analytical capability and exploit the range of earnings data sources and indicators available. The focus is on gaining the benefits of microdata analysis, especially by linking microdata from various sources, to maximise analytical power.

In response to the 'Review of statistics on distribution of earnings' (also called 'the Distribution of Earnings Review'), the Earnings Analysis branch developed a framework for labour cost statistics to identify the interactions, linkages and potential uses that can be made of the data. The branch published 'A framework for labour cost statistics' which draws together the needs of stakeholders to explain why the current set of data is collected. Requirements for data that are not met by the current set are also identified.

While the overall picture of earnings from the data sources is similar, there are important inconsistencies that arise because the data are collected through different channels and summarised using different methods.

This article reports on a study that draws together these various sources of UK earnings data for comparison at the macro level, to investigate their similarities and inconsistencies. The work highlights the need to bring sources together to enhance analytical capability and gain greater understanding of the coherence between data sources. As a result, one of the main objectives of the Earnings Analysis branch over the past year has been to link data sources at the micro level to provide a more comprehensive picture of labour costs. The findings of two linking projects are described here, along with their relevance for data collection and labour market analysis.

Earnings Analysis branch is also carrying out validation studies on individual datasets. These results are preliminary and so not reported here, but early indications are that even small changes in estimation methods can have a significant role in explaining the difference in, for example, low pay estimates. These results will be published later in 2007.

## Comparison of data sources

### Similarities

The main sources of earnings information collected by ONS are:

- the Annual Survey of Hours and Earnings (ASHE) (previously the New Earnings Survey (NES))<sup>1</sup>
- the Monthly Wages and Salaries Survey (MWSS)<sup>2</sup>
- the Labour Force Survey (LFS)<sup>3</sup>

ASHE and the MWSS produce earnings information as their primary focus, while the LFS considers earnings as part of a much broader survey. In addition, there are also various supplementary sources that contain information on earnings. However, these are only used to provide auxiliary information for analysis, because earnings are not comprehensively covered in these surveys and may therefore not be produced to the same standard.

Sources often provide overlapping information on earnings so they would be expected to produce similar results where this is the case. For example, most sources provide information on earnings at the individual level. However, the MWSS and the National Accounts provide information at the organisational and aggregate economy level respectively.

Data are frequently broken down by personal characteristics. Sex, age, occupation and full-time/part-time status are popular categories for all of the data sources. However, household surveys generally expand on this to provide richer personal information because the nature of the survey allows them to approach individuals directly. ASHE, being sourced from pay records, has much more limited personal information, while the MWSS is unable to collect any personal information because it is based at the organisational level.

Earnings sources also break down the data by company information, such as industry, location and company size. These data are equally covered by both employee and employer surveys. However, the latter is believed to be more reliable because it is collected directly from the employer. As such, any discrepancies between datasets tend to be resolved in favour of the employer-provided data.

### Inconsistencies

Despite similarities in the type of data that are collected, the different nature of the surveys will lead to inconsistencies in the results. In fact, inconsistencies can arise in the sources of data, the way that data are analysed, and the way the results of surveys are aggregated to produce estimates for the macro-economy. These three areas of inconsistency are now assessed in more detail, concentrating on the main data sources.

### Differences in the level of earnings: weekly earnings

The most striking inconsistencies are created between employer and household surveys. Employer surveys, such as ASHE, are considered more reliable because they gain access to precise wage figures through payroll records. Household-based surveys, such as the LFS, are less likely to be based on payroll records, and hence are more likely to be subject to a variety of recollection and measurement errors. In addition, proxy responses might be used to collect the data indirectly through other members of the household. This approach accounts for about 30 per cent of responses and may reduce the reliability of the results. There is also a fundamental difference across the two surveys in the information they collected on hours worked. As ASHE information is provided from employers, information will relate to contracted hours. The LFS asks respondents for information on the hours they have worked, which will not always be the same as their contracted hours. Hence ASHE and the LFS measure different concepts on the hours of work.

The NES, forerunner to ASHE, was seen as more reliable relative to the LFS because it used employer data retrieved from payroll records. However, it also suffered from limitations because it excluded individuals who were under the tax threshold. A large proportion of this category are likely to be part-time workers who earn a low hourly pay and will therefore lead to overestimated results. In addition, individuals who moved employers more frequently were more

likely to be missed. ASHE was designed with two supplementary samples to overcome these issues.

Mean and median weekly earnings information can be obtained from ASHE, the MWSS and the LFS. The figures are displayed in **Table 1** below along with some other ad hoc supplementary data sources. However, the majority of supplementary sources, such as the Family Resources Survey, cannot be quoted because the information is collected in bands rather than absolute figures. Where possible the median is used as the main measure because it is not biased by large outliers at the top of the distribution.

The LFS only publishes results for full-time employees and therefore comparable data have been provided for ASHE. It should be noted that this will increase the median figures because pay is generally lower for part-time employees. An additional issue for consideration is the different definitions used to categorise individuals as full or part time across the surveys. There is no agreed international definition as to the minimum number of hours in a week that constitute full-time or part-time work and the approach differs depending on the data source used. The LFS asks people to classify themselves as either full time or part time, based on their own perceptions. In ASHE, individuals are classified as part time if they work less than 30 hours per week.

The most significant results from the table are the consistently higher estimates of gross weekly earnings produced by ASHE, in comparison with the LFS, when considering full-time employment.

**Table 1**  
**Gross weekly earnings from various sources**

Source	Year	All Median	Males Median	Females Median	All Mean	Males Mean	Females Mean
ASHE FT <sup>1</sup>	2004	420	460	357	499	549	417
	2005	431	472	372	517	569	436
ASHE Total <sup>2</sup>	2004	346	435	259	414	514	309
	2005	350	440	268	423	526	320
LFS <sup>3</sup>	2004	380	420	323	451	438	374
	2005	385	423	335	465	508	334
MWSS <sup>4</sup>	2004	-	-	-	369	-	-
	2005	-	-	-	384	-	-
HMRC <sup>5</sup>	2003-04	-	-	-	393	-	-
EFS/FES <sup>6</sup>	2003-04	-	-	-	277	-	-
SEH <sup>7</sup>	2002-03	358	-	-	506	-	-

#### Notes:

- Not possible – no data available
- 1 ASHE excluding part-time workers.
- 2 ASHE including part-time workers.
- 3 LFS quarterly adjusted, using quarter 2.
- 4 AER – April 2004, whole economy.
- 5 Total annually earned income divided by 52.2.
- 6 One adult non-retired household.
- 7 England, annual income divided by 52.2.

The results indicate that the introduction of advanced surveying techniques to improve coverage of the low paid has not materially changed this picture. This difference has been widely noted, but there has been little work to explain it until now. This is discussed below.

A second important issue is the low estimate of earnings produced by Average Weekly Earnings (AWE). AWE is a monthly measure of average weekly earnings per employee and is calculated from the MWSS. A like-for-like comparison cannot be made with the LFS because AWE is based on both full-time and part-time employees. However, the difference between the two measures, if anything, will be wider if they covered the same sampling frame because of the nature of part-time earnings. Conclusions regarding the reasons for these low estimates are still to be reached; however, its survey design offers various potential explanations. For example, AWE calculates bonuses differently from ASHE, and a small role might be played by the fact that AWE adds weights to third and subsequent jobs.

HMRC is the only supplementary source in the table that provides comparable results with the main data sources. The low estimate provided by the Expenditure and Food Survey (EFS) is not directly comparable because it surveys total household earnings and so this figure is only based on households with one member. In addition, the Survey of English Housing (SEH) is likely to overestimate earnings, because it is based on total income, which includes things such as social security benefits and interest on savings.

Due to these comparability issues, the supplementary sources are not discussed in further detail.

**Table 2** and **Table 3** show ASHE and LFS earnings broken down by industry and occupation respectively.

ASHE earnings data broken down by industry still provide consistently higher estimates in comparison with the LFS. In addition to the comparability problems noted above, differences arise in the allocation of industries despite the fact that they use the same classification system (the Standard Industrial Classification). Differences may arise because the LFS is based on self-reporting. It asks employees to report the main activity of the organisation they work for so that the survey can use this information to classify their industry. The problem with this approach is that respondents may not always give an

**Table 2****Gross weekly earnings from ASHE and the LFS: by industry**

Mean, full-time employees	ASHE		LFS	
	2004	2005	2004	2005
All employees	499	517	451	465
Agriculture and fishing	386	398	294	334
Construction	506	527	456	453
Hotels and restaurants	319	325	360	362
Transport and communications	495	507	441	461
Banking, finance and Insurance	669	693	573	572
Public administration, education and health	491	516	453	468
All services	501	520	453	464

**Table 3****Gross weekly earnings from ASHE and the LFS: by occupation**

Mean, full-time employees	ASHE		LFS	
	2004	2005	2004	2005
All employees	499	517	451	465
Managers and senior officials	736	776	658	685
Professional occupations	670	708	642	650
Associate professional and technical	532	548	485	500
Administrative and secretarial	344	352	319	336
Skilled trades	421	432	382	390
Personal services	292	300	262	262
Sales and customer services	283	284	268	269
Process, plant and machine operatives	380	396	349	359
Elementary occupations	307	314	272	288

accurate description of their organisation's main activity. In contrast, ASHE uses employer information to assign all the employees of a company to the industry that is most prevalent within that company. For example, everyone who works in a local authority is assigned to the education sector as this is the main activity of most local authorities. For private sector companies, the industry classification is assigned to the area of activity that accounts for the largest share of the organisation's turnover.

Earnings data broken down by occupation, provided in Table 3, give a more consistent picture across the two surveys. In comparison with industry data, this might be because respondents in the LFS are more likely to describe their own job accurately (which is used to classify their occupation) than they are to give an accurate description of their organisation's main activity (which is used to classify the industry).

**Differences in the derived results: low pay**

ONS's low pay estimates are counts of the number of jobs paid below the NMW.

**Table 4** provides estimates of low pay for 1999–2005 from a number of different sources.

Before 2004, the National Statistic for the estimate of low pay was the central estimate, calculated by averaging NES and LFS

estimates. The central estimate was seen as a way of balancing the limitations of both the LFS and the NES, as noted above, to provide more reliable results.

In 2004, the ASHE survey was introduced to replace the NES, to improve on the representation of the low paid using two innovations. First, the sample was supplemented by including employees in businesses outside the PAYE system, that is, by selecting VAT-only businesses from ONS's Inter-Departmental Business Register (IDBR). Second, the ASHE/NES sample was taken in February, which excluded individuals from the analysis who moved during the time the sample was selected and the survey date. ASHE improved coverage of these mobile workers (who are more likely to be low paid) by taking a later sample to catch the leavers and joiners. Weighting and imputation also improved ASHE outputs relative to the NES. As a result, the ASHE estimate of low pay became the National Statistic, with the LFS running alongside as a source of complementary information on personal characteristics of the low paid.

Since 2004 the LFS has seen improvements of its own. For household surveys, stated hourly rates are more reliable than hourly rates derived from total earnings and hours worked (the derived rate), due to reporting errors. This can

**Table 4**  
**Jobs paid below the NMW for individuals aged 18 and over: by estimate method**

Percentages					
	LFS 2002 method	LFS 2005 method	ASHE	NES	Central
1999	2.2	-	-	2.1	2.2
2000	1.0	-	-	1.0	1.0
2001	1.1	-	-	0.9	1.0
2002	1.5	-	-	1.2	1.4
2003	1.0	-	-	1.1	1.0
2004	1.2	0.7	1.0	1.0	1.1
2005	1.7	1.1	1.2	-	-

**Note:**

- Not possible – methodology not applicable

be clearly seen in the LFS data, because derived rates of hourly pay have a much wider distribution than the stated rates and are less plausible. Regardless of this, stated rates were only collected for employees in their first job before 2004, which meant any measure of low pay for second jobs had to be calculated using the derived rate.

In 2004, the LFS introduced an additional question to include stated hourly pay for second jobs. In Table 4, the second column shows the LFS estimates calculated using the amended methodology. It shows that the LFS and ASHE estimates have now converged to provide more consistent results and that the new LFS figures actually report lower estimates of the low paid in comparison with ASHE.

There has been some preliminary work, to be published this year, indicating that the remaining differences between the new LFS methodology and ASHE are largely down to the measure of hourly pay used. When ASHE is calculated using the same methodology as the LFS, that is, using a combination of the stated and derived rate, the difference between the two measures is minimal; the investigation therefore reconciles the difference between ASHE and the LFS low pay estimate. Nevertheless, these are only preliminary results, so further work is needed in this field to validate these claims.

### Differences in the aggregated results: the macro-economy

National Accounts figures are currently derived independently from labour market statistics and an inconsistency was therefore expected. The Productivity Economics branch in ONS has investigated the inconsistency between the National Accounts whole economy wage bill and other sources of labour market statistics.

This work involved comparing the derived whole economy wage bill from the variety of sources above to the National Accounts figure. Key results were:

- ASHE-derived figures consistently overestimate the wage bill compared with the National Accounts figure
- the LFS- and AWER-derived figures consistently underestimate the wage bill compared with the National Accounts figure
- the additional supplementary sample that was introduced in the ASHE survey in 2003 brings ASHE and National Accounts figure closer, although it is still too early to determine if the same impact will occur for the subsequent years

A similar result was found by looking at quarterly figures from the LFS and AWER.

Hence, the National Accounts figure lies between the household and employer survey estimates. This may be because, while the raw earnings data in National Accounts are similar to ASHE as both are derived from employers, the aggregate wage bill is weighted to LFS jobs.

### Linking data sources

Given that there are inconsistencies between the data sources on earnings, it has become apparent that linking these sources could prove fruitful in a number of ways. Linking earnings data sources at the macro and micro level will help to validate the different sources and act as a quality control check. In addition, it will also provide a larger dataset to broaden analysis. Data linking between household and employer surveys is seen as particularly important because their strengths lie in different areas. However, some of these

very inconsistencies mean that linkages may be hard to establish. A feasibility study was therefore carried to assess whether data linking could take place.

### Feasibility study

The first part of the feasibility study is outlined in 'Linking earnings data: Methods'. It looks into the different methodological approaches that can be used to link datasets. The four main methods are:

- direct record linkage – this is the optimum form of linking and can take place when there is a unique error-free identifier attached to each record in each dataset from the same population
- probabilistic record linkage – this is an imperfect method used when there are errors in identifiers, such as spelling mistakes, but given they come from the same population, they can be matched with a given degree of certainty
- data fusion/statistical matching – datasets from different populations or with no identifiers, for example, can use this approach to achieve a merged dataset, based on 'representative' individuals with similar characteristics, rather than linking specific individuals, and
- cell group linking – this method creates matching groups of representative individuals. It is therefore a generalisation of data fusion based on groups rather than individuals

These methods were then assessed against a number of projects in a feasibility study entitled 'Linking earnings data: Benefits of linking and possible projects', in order to establish the benefits and practicalities of merging specific datasets. All the projects were supported by the main users of earnings data, with the view that they would help develop the analytical capabilities in this area.

The linking projects fall into three broad categories:

- linking supportive information on companies to earnings information, to allow analysis of company characteristics and attitudes towards pay
- linking employer- and employee-provided earnings information. This would combine the more reliable data from employers with the more extensive coverage of personal characteristics from the employee surveys, and

- linking supportive information on personal characteristics to earnings information. This is a similar approach to the one above, but uses other supportive information, for example the UK Census, to link to employer surveys

The feasibility of the projects was then assessed against criteria such as the extent of benefits, ease of obtaining data, known difficulties and methodological opportunities.

The ASHE-IDBR linking project under the first category above was chosen as a study into direct linkage. This project proved enlightening because it allowed analysis of company attitudes towards the NMW. The IDBR could be accessed with immediate effect and provided a good likelihood of matching against the MWSS as well.

The ASHE-LFS linking project under the second category above was carried out to investigate data fusion techniques. Data availability for this project allowed the work to be initiated immediately and the methodological approaches developed in the investigation can be utilised in other areas in the future.

The third project relies on the experience gained in the first two projects. With the success of these, the Earnings Analysis team is now evaluating a wider range of linking projects for feasibility in this area (see below). The next sections describe the projects and main findings.

### Linking earnings and firms: the ASHE-IDBR project

The purpose of the investigation was to provide a linked dataset by merging employer information from ASHE with information on companies from the IDBR. The resulting dataset could then be used to examine how companies set wage policies at the lower end of their pay scale, given changes in the minimum wage. This is an important area of investigation because previous analysis has focused on pay from the viewpoint of the worker and ignored the behaviours of company policies. The initial investigation exploited a variable unique among large scale datasets to examine the changing wage for a job. The second part of the investigation linked employer and employee data together to look more broadly at how and if companies' wage policies respond to changes in the NMW.

The investigation concluded that there are strong company effects that influence wage-setting policies. Rather than setting wages equal to the marginal cost of labour, companies are also employing a relatively

simple rule-of-thumb approach for wages around the NMW. The main findings of the investigation are:

- there is evidence that, as the NMW increases, the salaries of all low-paid individuals increase by much the same amount regardless of their distance from the minimum wage. This effect is known as relocation, and is counter to compression, which occurs when an increase in the minimum wage has no effect on wages above the new level, but raises those below it just up to the new NMW
- the investigation introduces the concept of the Company Minimum Wage (CMW), that is, the minimum wage paid by a particular company in a particular year
- there is evidence to suggest that these CMWs are set relative to 'focus' points, such as £5.00, £5.50, despite the fact that the NMW does not reflect these round numbers. This suggests firms have some flexibility in the way they set wages and they are not wholly driven by the NMW
- this preponderance of 'focus points' continues up the wage distribution, and is observable up to 80 per cent over the minimum wage
- there is evidence that companies prefer to maintain wage differentials relative to general labour market conditions. The NMW contributes to the absolute level of wages, but is not the only or the dominant factor
- a brief examination of industry differences suggested that the numbers might differ but that the story remains broadly the same

More generally, the linked employer-employee dataset used in this investigation provides a substantial research resource. Although the linked ASHE-IDBR dataset is not a full rectangular one (all employees linked to all firms), there has been considerable interest in the academic community and government. The linked dataset is in the process of being documented and released to the research community through ONS's Virtual Microdata Laboratory (VML).

### Linking employer and employee data: the ASHE-LFS project

This project links the ASHE and LFS datasets. ASHE and the LFS represent the main sources of earnings information in the UK and form the basis for most micro- and

macro-level analysis of the labour market. Linking the two datasets is an important development for earnings analysis because the two sources do not provide a single view of the labour market. The two surveys are designed for different purposes and use different surveying techniques (and sometimes use different concepts and definitions) that lead to inconsistencies, as outlined above.

Unfortunately, there is little overlap between the two surveys (theoretically only 1 per cent of the LFS sample should be found in ASHE) and they do not share a common direct identifier. As a result, direct record linking cannot be employed to merge the two datasets. The investigation employs a grouped cell method to create a linked dataset that contains properties of both data sources, which can still be analysed relatively robustly. The resulting dataset has two aims:

- to test statistical properties of the combined variable set so that more detailed models may be drawn, and
- to analyse the dataset in its own right

The central result of the investigation indicates that the ASHE and LFS datasets are more consistent with each other than previously thought. More specifically, gross weekly pay stands out as being particularly well-related across the two surveys and across a large proportion of the distribution. This is a welcome result because access to ASHE is limited for non-government-based users and therefore the LFS is used a good deal for academic research. Researchers should therefore be more confident using the LFS as an unbiased estimate of ASHE and the fact that results are not affected by data collection methods. Any inconsistencies that do exist between the two datasets tend to be focused around the top and bottom end of the distribution. This is based upon looking at comparable groups. Differences appear in the broad aggregates because of differences in coverage; in particular the smaller sample size of the LFS appears to lead to poor representation of some population groups.

To some extent, this finding eliminates the need for further evaluation of earnings using a linked ASHE-LFS dataset; the LFS alone may be good enough to examine characteristic breakdowns that are not averaged on ASHE. However, further analysis of ASHE earnings data using LFS breakdowns suggests this might not be the case for certain characteristics. Ethnicity, for example, is based on small sample

sizes in the LFS and it is not possible to make reliable inferences from this data. Therefore estimates of earning distributions by personal characteristics are more robust for categories that have larger sample sizes across the whole range of potential values, such as skill.

## The way forward

### The ASHE-IDBR and ASHE-LFS linked datasets

For both the linked datasets, the plan is to release them to researchers. ASHE-LFS work shows this link is less important, but releasing the datasets to researchers will enable the conclusions presented here to be tested more thoroughly by external experts. There is considerable interest in the ASHE-IDBR linked dataset, as it can be used directly in conjunction with other ONS business data (already linked to the IDBR), and some work has already been carried out in this area.

Both datasets will be released in ONS's VML. This is an extremely secure research facility available at ONS sites and other selected locations, which is designed to give maximum access to data consistent with a strict confidentiality regime.

The linked ASHE-LFS dataset does not pose any particular confidentiality problems, as data items are created based upon statistical characteristics and do not increase the identification risk for any individual in either survey. The ASHE-IDBR link does increase the potential for re-identification of firms or workers, but this is manageable within the existing VML framework. No additional confidentiality concerns are therefore likely to arise as a result of releasing these datasets.

### Further linking

Following the success of the first two linking projects, additional projects are now being considered. The first to be pursued is a feasibility study into the ways that ASHE, the Census, the National Pupil Database and the LFS might be usefully combined. For example, ASHE and Census data linking could provide further breakdowns of ASHE data by personal characteristics. Since individuals in ASHE should also appear in the Census, direct linkage may be possible. However, difficulties in identifying matching individuals may mean data fusion techniques need to be employed.

JUVOS is a longitudinal database of statistics on claimant unemployment and merging it with ASHE will help to analyse the link between earnings and employment. A common identifier and sample should make this a more robust link.

### Other earnings projects

Away from data linking, a project on the earnings of the self-employed is being pursued. This area has received little coverage in the past and is limited by a lack of clarity over its definition. However, if this problem and others can be overcome, then the project should help to create a more complete picture of earnings in the UK.

### Conclusion

This article has reviewed a number of projects carried out over the past eighteen months. These have produced a number of important results, both in relation to the datasets involved and in the understanding of the UK labour market.

Although many of the results from the work programme focus on low pay, the methods and analyses have a wider importance. These are being expanded upon in the current work programme to study the relationship between earnings and self-employment, unemployment, educational attainment, and so on. The results of these will be discussed in the 2008 'Earnings: Summary of sources and developments' publication.

#### CONTACT

✉ [elmr@ons.gsi.gov.uk](mailto:elmr@ons.gsi.gov.uk)

#### REFERENCES

- Bird (2002) 'Review of distribution of earnings statistics (National Statistics quality review series, report no. 14)' at [www.statistics.gov.uk/methods\\_quality/quality\\_review/downloads/DOER\\_Final\\_Report.doc](http://www.statistics.gov.uk/methods_quality/quality_review/downloads/DOER_Final_Report.doc)
- Lam K, Ormerod C (2005) 'Linking earnings data: Methods', mimeo, Office for National Statistics

#### TECHNICAL NOTES

- ASHE is a survey of employers requesting individual level information about their employees. The survey is a 1 per cent sample of employees on the Pay As You Earn (PAYE) register. Employers are asked to provide information on the hours and earnings of their employees. ASHE is a new survey that has been developed to replace the New Earnings Survey (NES) since 2004 to include improvements in the coverage of employees and to the weighting of earnings estimates. The data variables collected remain broadly the same, although an improved questionnaire was introduced for the 2005 survey.
- The Monthly Wages and Salaries Survey (MWSS) is a questionnaire sent to a sample of businesses on the ONS business register. This sample usually comprises 8,600 companies which are then split into public and private sector and then by industrial classification. The companies are then split into bands dependent on the number of employees. All companies in the largest bands are sampled and a random sample is taken from companies in the smallest bands. Companies with fewer than 20 employees are not included. Information is collected on the total wages and salaries bill of employers; details about individuals are not collected. The purpose of MWSS is to provide the base data for the calculation of short term indices of earnings.
- The Labour Force Survey (LFS) is an annual survey of households in the UK. The survey collects information about the household members and their characteristics. As well as earnings information, the survey includes questions on household members' jobs, employment pattern, sickness, benefit entitlements and health.

Lam K, Ormerod C, Ritchie F, Vaze P (2005) 'Do company wage policies persist in the face of minimum wages?', *Labour Market Trends* 114(3), pp69–82 and at [www.statistics.gov.uk/articles/labour\\_market\\_trends/minimum\\_wage.pdf](http://www.statistics.gov.uk/articles/labour_market_trends/minimum_wage.pdf)

Lau E, Porcu S (2005) 'National Accounts' Compensation of Employees: a comparison with Labour Market Statistics', mimeo, Office for National Statistics

Ormerod C (2005) 'LFS low pay estimates: Method of dealing with second jobs', *Labour Market Trends* 114(7), pp233–9 and at [www.statistics.gov.uk/downloads/theme\\_labour/LMT\\_July06.pdf](http://www.statistics.gov.uk/downloads/theme_labour/LMT_July06.pdf)

Ormerod C (2005b) 'Linking earnings data: Inconsistencies and similarities between different sources', mimeo, Office for National Statistics

Ormerod C (2006) 'A framework for labour cost statistics', *Labour Market Trends* 114(6), pp203–11 and at [www.statistics.gov.uk/downloads/theme\\_labour/LMT\\_Jun06.pdf](http://www.statistics.gov.uk/downloads/theme_labour/LMT_Jun06.pdf)

Ormerod C (2006) 'Earnings data: A brief guide to sources and outputs', *Labour Market Trends*, 114(11), pp389–96 and at [www.statistics.gov.uk/downloads/theme\\_labour/LMT\\_Nov06.pdf](http://www.statistics.gov.uk/downloads/theme_labour/LMT_Nov06.pdf)

Ormerod C (2006) 'Earnings data: Sources and outputs' at [www.statistics.gov.uk/articles/nojournal/1639.pdf](http://www.statistics.gov.uk/articles/nojournal/1639.pdf)

Ormerod C (2006) 'Linking ASHE and LFS: Can the main earnings sources be reconciled?', mimeo, Office for National Statistics

Ormerod C (2006) 'Linking earnings data: Benefits of linking and possible projects', mimeo, Office for National Statistics