The Average Earnings Index and Average Weekly Earnings

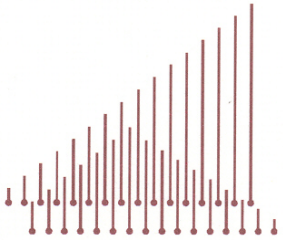
Report submitted to the National Statistician by

Dr M.R. Weale,
National Institute of Economic and Social Research

2nd June 2008

CONTENTS

Letter from Martin Weale to the National Statistician



National Institute
of Economic and
Social Research

2 Dean Trench Street
Smith Square
London SW1P 3HE
Telephone +44 (0) 20 7222 7665
Facsimile +44 (0) 20 7654 1900
e-mail enquiries@niesr.ac.uk
http://www.niesr.ac.uk

2nd June 2008

Ms K. Dunnell,
National Statistician,
Office for National Statistics,
Newport,
Wales.

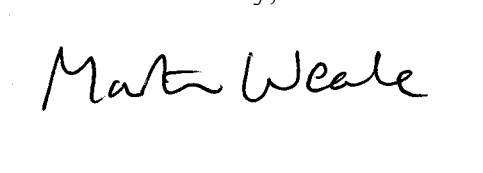
Dear Karen,

At the start of February 2008 you asked me to review aspects of the experimental series of Average Weekly Earnings produced by the ONS and I now submit the following report. I recognize that the Review has taken longer than initially hoped. This is because the issues involved were complex and it took some time to come to grips with them despite the fact that the documentation made available by your colleagues was of high quality. Nevertheless, I hope the Report will prove useful to ONS in bringing the Average Weekly Earnings series to a state where it can be put forward for assessment as a National Statistic.

This work would not have been possible without the substantial input of Derek Bird, Peter Brodie, Harry Duff, Neil Parkin and Paul Smith. Harry Duff and Neil Parkin have in particular devoted a substantial amount of time to explaining the calculation of Average Weekly Earnings and the way in which this differs from the calculation of the Average Earnings Index. Peter Brodie provided substantial input into section 5.3, Harry Duff provided overall background and the material for Chapter 7, Neil Parkin carried out the calculations needed for Chapter 4 and Paul Smith provided substantial input for section 5.2 and Appendix 3.2 and Appendix 4.1. Peter Brodie and Martin Brand worked on issues raised by Appendix 5.

Very helpful comments on and corrections to earlier drafts have been provided by these ONS staff, by Colin Mowl and also by Chris Kelly of HM Treasury and Neal Hatch of the Bank of England. Of course responsibility for both the text of this report and any errors and omissions is entirely mine.

Yours sincerely,



Chapter 1. Summary of Findings and Recommendations

1.1 Terms of Reference

The terms of reference for this Review are to investigate:

1. whether the ONS's proposed methodology for Average Weekly Earnings (AWE) is fit for purpose for the principal measure of short-term changes in earnings;
2. whether the reconciliation between AWE and the Average Earnings Index (AEI) provides a satisfactory explanation of the statistical and economic differences between the two measures;
3. whether AWE is a better measure of short-term changes in earnings than the AEI;
4. which AWE measure (fixed weights or current weights) should be used in which circumstance;
5. the nature and extent of supporting information required to supplement AWE;
6. whether, after a transitional period, the AEI can be withdrawn;
7. whether the current sample structure is appropriate, particularly with reference to the financial services sector.

In interpreting the meaning of "short-term changes in earnings" it has been assumed that these relate to the growth in earnings measured by comparing the earnings indicator (AEI or AWE) averaged over the last three months with the average for the same period one year earlier. The growth rate thus defined is referred to as the annual growth rate.

1.2 Findings

These findings are reported in the order in which they emerge from this report. Their connections with the terms of reference are indicated.

The ONS has instigated a considerable development programme in recent years, partly in response to the review of the AEI carried out by Sir Andrew Turnbull and Mr Mervyn King (Turnbull and King, 1999) in which I was heavily involved, to improve their short-term earnings indicators. The AWE is a product of this work and has been developed as an improvement on the AEI and on the assumption that it will eventually replace the AEI. There are good reasons for believing it to be superior to AEI as a wage indicator:

- its methodology is designed to enable it to measure the average weekly wage per person employed in the economy;
- unlike AEI, AWE is responsive to the effects of shifts in the composition of employment in the economy on the average wage;
- AWE handles non-response in a manner which is probably better than AEI; for an indicator of the level of wages
- AWE makes an explicit adjustment for the smallest businesses in the economy, which are excluded from the survey that produces the information on which AWE and the AEI are based, by using auxiliary information from the large scale annual survey of earnings. Conversely, in

the AEI small businesses' pay growth is assumed to be the same as that of larger businesses.

These are some of the factors that make AWE a strong candidate to replace the AEI, but before it does it is important to be clear that all aspects of the methodology are at least as good as that of the AEI. This is because there are sometimes substantial differences between the growth rates in earnings computed from the AEI and the estimates produced by the experimental measure of AWE. Nevertheless, the two indicators are much closer to each other than is either to a range of possible alternative indicators of earnings growth. Thus it is not possible to choose between them by means of validation against a generally accepted third measure.

Nevertheless there is a question whether all of the methodological divergences of AWE from AEI are justifiable if the aim is to produce a measure of short-term changes to earnings (Term of Reference 1). At this point it should be noted that some users find information on the level of earnings is helpful and that this creates some tension with the aim of providing a good indicator of short-term earnings growth. The implications of this are discussed in section 5.6 and have important bearings on the recommendations of this report.

The ONS has concluded a significant and careful piece of detailed methodological research to explain the reasons for the differences between the indicators. This, and the work undertaken as part of this review, shows there are two sources of the differences. First of all, AWE is designed to be different. The AWE is a genuine measure of average earnings whereas the AEI is not because changes in the AEI do not reflect the impact on the true arithmetic mean of changes in the composition of employment in the economy.

Secondly, differences arise because in designing the AWE the ONS made a number of judgements, influenced in part by research on other indices which it produces, which are either implicitly or explicitly different from those underlying the AEI - for example, in the treatment of missing data/sample rotation and outliers. This work was taken into account by ONS when it designed AWE in response to recommendations in the Turnbull-King report. The design work and delivery were of a high standard but, given the wide scope of statistics produced within ONS, some of the judgements, while reasonable, may not have been optimal if the focus is on short-term changes in earnings as defined above. Nevertheless, it is possible to do further research to establish whether the approaches it has adopted to handling missing observations/sample rotation and outliers are in fact better than those used in the AEI or whether there are other possibilities which might be superior to both.

Finding 1 Term of Reference 1 *See Chapter 2*
The AWE offers in principle a better measure than the AEI of changes in the average wage in the economy. However, subsequent findings indicate that more work is desirable to ensure that it is also in practice a better measure.

Finding 2 *See Chapter 3*
There is a range of indicators of wage movements available as alternatives to AWE and AEI. Some of the indicators explored are slightly more closely related to the AEI than AWE while for other indicators the opposite is true. There is some concern that

the AEI tracks the movement of wages and salaries per employee in the national accounts better than does AWE. But the most important point to observe is that, when looking at monthly or quarterly growth in earnings, the AEI and AWE are much more closely related to each other than they are to any of the other measures. Thus a general conclusion which follows from this section is that other data series do not provide any solid guidance as to whether AWE should be preferred to the AEI or vice versa.

Finding 3 Terms of Reference 2 and 4. *See Chapter 4*
ONS's decomposition of the differences between AWE and the AEI is coherent and provides an extensive assessment of the impacts of different features of the construction of the series on the differences between them. Nevertheless a slightly different decomposition has been presented in this chapter. A particular feature of this is that it separates the impact of changes in employment structure on the average wage into two components- those which arise from the effect of changes and those which arise because those changes have a subsequent impact on the way in which wages in individual industries are aggregated to produce an average. This decomposition makes it possible to identify the changes to the average wage which would occur in the absence of employment composition effects. This measure rather than a fixed-weight index should be used in order to understand the effects of wage changes within industries leaving aside the impact of changes in the mix of employment across industries (Recommendation 4).

Finding 4. Term of Reference 1 and 3 *See Chapter 5*
As an indicator of the level of wages the AWE is close to fit for purpose. However, if the primary aim is to identify short-term changes to wages, further work is needed to establish whether the AWE methodology is better than the AEI methodology. (Recommendations 1 and 2). ONS has the clear capacity to carry out the required work effectively. There are also modifications which can be made without further research which should make the AWE more satisfactory (Recommendation 3). These have implications for the form in which AWE is presented and in turn imply that further work is needed to provide an indicator which meets the needs for both estimates of growth in earnings and of the level of earnings. (Recommendation 5).

Terms of Reference 5 and 6 are addressed in Chapter 8 and Recommendation 8. Term of Reference 7 is addressed in Chapter 6 and Recommendation 6.

1.3 Recommendations

Given the importance of the AEI and thus of a possible replacement, the ONS should be able to assure users of the data that, in making decisions on methodology, it has fully explored the possibilities. This report therefore recommends:

Recommendation 1. (page 44). The role of matched pairs in dealing with missing observations and sample rotation
AWE should not become a National Statistic until further work has been carried out on the possible use of matched pairs. This work needs to compare the use of matched pairs or a combination of imputation and matched pairs with the existing AWE methodology to see which produces more reliable estimates of annual growth rates.

Similarly, it needs to be established that the treatment of outliers is appropriate, given that the primary aim of the measure is to provide an indicator of wage growth

Recommendation 2. (page 45). The treatment of outliers
AWE should not become a National Statistic with the treatment of outliers present in the experimental series until it is clear that this is superior to the alternative which has been demonstrated to be workable in the construction of the AEI (but which would need development to be applied to AWE).

The report outlines the work to be done so that ONS can feel confident it has identified the best of the possible candidates.

It was also found that changes in employment as a result of firms being reclassified from one industry to another were, from time to time, a source of disruption to the AWE. Many of these changes arise from improvements to the quality of estimates in the Interdepartmental Business Register (IDBR) and it is not desirable that they should be a source of disturbance.

Recommendation 3. (page 46). Employment Classification
The AWE series should be linked across substantial employment reclassifications so that the effects of reclassifications themselves do not affect the estimated rate of growth of the average wage.

The review was asked to investigate whether there should also be a fixed weight version of the AWE (this series would still reflect any methodological differences from the AEI which still remain following the work carried out under recommendations 1 and 2 but it would be computed using fixed employment weights). Users may well wish to know how far the changes in average earnings arise because of changes in earnings in individual industries and how far they arise because of factors such as changes to the employment mix between industries. However, this need is not best met by a fixed weight index whose employment weights are revised only intermittently. Instead it can be met by separating out the effects of wage changes from the effects of employment changes in the AWE. A useful distinction can be made between the use of chain-linked indices and fixed-weight indices. The former are generally preferred to the latter and have superseded them in other contexts such as the national accounts. The proposal here is therefore for a chain-linked index rather than a fixed-weight index.

Recommendation 4. (page 46). The Impact of Changes to Wages in Individual Industries
The ONS should publish, along with the main AWE results, figures showing the change to average earnings arising solely from changes to average wages by industry, and excluding any employment mix effect.

Recommendation 3 and, depending on the outcome of the work proposed, Recommendations 1 and 2 taken on their own would imply that it would not be possible to provide a meaningful estimate of earnings in pounds per week. This would not be a problem if the sole aim were to produce an indicator of short-term changes in earnings. However, given the user interest in the level of average earnings

the ONS needs to do work on producing coherent and satisfactory measures of the level and growth of earnings.

Recommendation 5. (page 49). Short-term Changes and the Level of Earnings
ONS should research means of producing AWE so that it serves primarily as a short-term indicator of earnings growth but also delivers a coherent estimate of the level of wages.

The review was asked to explore whether the sample structure was appropriate given the changes to the economy, and in particular, the changes to the financial sector since it was last revised. ONS has to revise the sample in 2009 in any case because the industrial classification is due to change. It is doubtful that the benefits of carrying out an extra revision ahead of this would outweigh the costs. However, ONS has proposed an interim increase in the sample size of the financial sector

Recommendation 6. (page 51). The Financial Sector
As proposed by ONS, the sample size in the financial sector should be raised by 100 firms as soon as possible. The sample structure should be reassessed in 2009 in the light of the new Standard Industrial Classification.

During the course of the review it became clear that the software used in the production of the AEI ensures that the statisticians involved have the capacity to analyze the factors causing change in the index in very great detail. It is easy for them to identify the effects of individual respondents on the aggregate series and thus straightforward for them to review sources of potential disturbance. The software used in the production of AWE does not offer the same functionality.

Recommendation 7. (page 53). Software
Before AWE becomes a National Statistic, ONS should ensure that the software used to produce it has the same functionality as the AEI software with respect to revisions, changes and flexibility and, given the importance of wage statistics, that quality assurance of the programmes and procedures be carried out by a competent expert who was not involved with their development. The work should be done with an awareness of the needs of key users to understand what lies behind the movement in AWE.

After the work proposed under Recommendations 1 and 2 and the implementation of Recommendation 3, the ONS will be confident that differences from AEI arise because the AWE is inherently better. Given this, there seems little point in continuing to produce the AEI. However, both statistics should be retained for a period of twelve months to allow users to adjust to the fact that change is coming.

Recommendation 8 (page 55). Replacement of AEI by AWE
It is recommended that ONS replace the AEI with AWE once AWE has been given the status of a National Statistic but that both be published as National Statistics for a period of twelve months. Assuming that the parallel run does not produce any further issues, the AEI should then be discontinued. AWE data should also be made available back to at least 2001 and further back if possible.

Chapter 2. The Average Earnings Index and Average Weekly Earnings

2.1 Introduction

There are two main measures of short term earnings growth, the Average Earnings Index (AEI) and Average Weekly Earnings (AWE). The AEI is the current National Statistic and is widely used as an indicator of short-term movements in earnings. Some perceived problems with this indicator in 1998 led to the Turnbull-King Review (Turnbull and King, 1999). The review noted that the concept underlying the average earnings index is not that of an average wage per person employed because a shift in employment from a low-paying industry to a high-paying industry or *vice versa* would not affect the index. It therefore recommended that the "ONS investigates the production of an index which reflects more closely movements in true average earnings" (recommendation 31). The AWE is the output of the work which was set in train after the Turnbull-King Review.

The Review also noted that there might also be a role for an index of wage costs, similar in form to the Retail Price Index. Such an index would look, effectively, at the wage rates for specific types of job and aggregate the movements; it would not be intended to reflect to any extent the impact of changes in the composition of employment on the wage bill.

The ONS explored both the definitive recommendation and the suggestion of a wage cost index. The measure of average earnings could be produced from the same Monthly Wages and Salaries Survey used to provide the data for the AEI, while an index of wage costs would require quite a different data source. The ONS concluded that the production of an index of wage costs was not practical but developed the recommendation that a measure of the average weekly wage should be produced. This led to the production of AWE, which was released as an experimental statistic in August 2005.

At the time of the launch of AWE as an experimental series, it was recognised that further development work was needed (for example on the treatment of outliers and the treatment of missing observations) before AWE could be launched as a National Statistic. This development work has recently been completed and is the subject of this review. In conjunction with this development work the ONS has also undertaken a programme of work (Parkin, 2008) to examine the differences between the AEI and AWE. This reconciliation work has also been examined in this review.

There are inevitably differences between the growth in earnings shown by the AEI and the growth shown by AWE. This is not in itself a cause for concern. While both series are calculated from the same data set, different calculations should be expected to lead to different answers. Despite the fact that it does not represent average earnings precisely, it is fair to say that the Bank of England and HM Treasury were reasonably happy with the existing index. But the existence of two indicators based on the same survey that, at times, give different messages, creates uncertainty for users.

This review can be seen as both an extension of the work by Parkin (2008) and as assuring the quality of it.

The terms of reference for the review are to investigate:

1. whether the ONS's proposed methodology for AWE is fit for purpose for the principal measure of short-term changes in earnings;
2. whether the reconciliation between AWE and the Average Earnings Index (AEI) provides a satisfactory explanation of the statistical and economic differences between the two measures;
3. whether AWE is a better measure of short-term changes in earnings than the AEI;
4. which AWE measure (fixed weights or current weights) should be used in which circumstance;
5. the nature and extent of supporting information required to supplement AWE;
6. whether, after a transitional period, the AEI can be withdrawn;
7. whether the current sample structure is appropriate, particularly with reference to the financial services sector.

At the outset, it should be stated that there is a clear logical basis for replacing the AEI with AWE; the latter is a measure of the average wage in the economy, whereas the former does not reflect the impact of changes in employment patterns. Nevertheless, in constructing AWE, ONS statisticians have looked afresh at many of the assumptions which inevitably have to be made in moving from survey data to aggregate statistics. An overriding issue is therefore whether these assumptions, where they differ from those of the AEI are likely to be improvements.

In answering this question some thought has to be given to the meaning of improvements and it depends, at least to some extent on the focus of users. Both producers and users of data need to bear in mind that an indicator designed to produce the most reliable indicator of the level of a variable, given an underlying survey, is likely to be different from an indicator designed to give the most reliable indicator of the growth rate. This is despite the fact that, if there were no sampling errors, the level and the growth rate would be arithmetically consistent. This point is illustrated in detail in Appendix 3.1. Thus a change which might be an improvement if the sole aim were to produce an indicator of the level of average wages could be a deterioration if the main focus were on indicating the rate of change.

As noted above, while the Terms of Reference mention only short-term changes to wages, users are in fact interested also in the level of wages. This means that ONS may well have to use its judgement to trade off performance in terms of levels as against performance in terms of short-term changes.

It has not been possible to examine fully all of the issues this point gives rise to in the time available for the preparation of this report; where necessary, therefore, suggestions are made for ways in which these issues might be investigated further and how ONS might reach conclusions on the balance between the quality of an indicator of the level of wages and the quality of an indicator of their growth.

This chapter describes the Monthly Wages and Salaries Survey (MWSS) and the compilation of the two measures. It then provides the reader with a summary of the differences between the two measures.

2.2 The Monthly Wages and Salaries Survey

The data source for both AEI and AWE is the MWSS. This is a monthly survey of 8,500 businesses with 20 or more employees. The sample is a simple stratified design with four employment bands (Band 1 20-99 employees, Band 2 100-499 employees, Band 3 500-999 employees and Band 4 all businesses with 1000 or more employees) coupled with a higher-level stratification (two digit standard industrial classification (SIC) by Public/Private Sector giving 116 categories). The term firm is used to refer to each reporting unit although they may not be separate legal entities. The term industry is used to refer to each of these higher-level strata although for each two-digit industry for which there are both public and private sector employees the two different sectors are treated as separate strata. The sampling frame is provided by the Interdepartmental Business Register (IDBR).

Only the largest size band is surveyed fully while the smaller strata have appropriate sampling fractions optimised for production of the AEI by minimising the variance of the month on month growth (last optimised in May 2003). Note that firms employing fewer than 20 employees are not surveyed.

The Turnbull-King report recommended that the survey questionnaire make clear which entries should include bonuses (recommendation 13). This recommendation has been followed and the questionnaire now in use is much clearer than that used at the time of the Turnbull-King review.

Neither measure takes direct account of monthly information on the movements in wages paid by small businesses, although the AWE does include an adjustment to mitigate this effect (see section 2.3). It has been suggested that movements in wage rates in very small firms are likely to be more sensitive to the state of demand than are movements in larger firms. The argument is that employment in these firms expands more rapidly than in the economy as a whole in cyclical upswings and falls more rapidly in downswings. Thus leaving small firms out of the survey may lead to biases. On the other hand, if policy has been set in the past on the basis of wage growth statistics whose fluctuations around the mean have been damped and the nature of the damping does not change, then the omission of very small firms should not be seen as a problem.

At some point the ONS may wish to investigate whether there is a case for including very small firms strong enough to outweigh the costs to both ONS and the providers of data which arise either because the sample has to be expanded or because the proportions of larger firms sampled has to decline. But that is obviously an issue outside the core of this report which compares two different ways of using the MWSS as it actually is.

2.3 The AEI and the AWE

As noted above the AEI combines estimates of the average wage by industry using fixed employment weights. The average figure that it implies for the whole economy

does not, as a consequence, take any account of effects arising from changes in the mix of employment between industries.

Over and above the effects of these differences in aggregating the industry results, there are a number of other methodological differences between the two series. This section summarises those key differences.

Imputation and Matched Pairs

In the AEI no explicit imputation is made for firms which do not respond to the survey. Instead the growth in the index is calculated from matched pairs, so that, when the growth rate is calculated between month *i-1* and month *i*, each firm is included in the sample for month *i* only if it is also present in the sample for month *i-1*. Thus a firm which enters the sample in month *i-1* will be used in the calculation of the growth rate of the index between *i-1* and *i* (provided it responds in both periods) but not in the calculation of the growth rate between *i-2* and *i-1*. The implication of this is that firms which do not respond are assumed to have the same growth rate in their average wages as is observed for the responding firms in the stratum. In AWE a simple imputation methodology is employed using recent data where available and appropriate but allowing weighting to compensate for missing data otherwise.

ONS has expressed concern that the use of matched pairs, as in the AEI, results in "drift". If wage growth in firms entering and leaving the sample is different from those which stay in the sample, then a bias may appear in the growth rate so that the index drifts away from the value which would be calculated if the average were calculated afresh in each period. Similarly, and perhaps more importantly, if a firm does not respond to the survey in any single month for some reason or other, it is assumed that, in the month in which it is absent, its average wage grows in the same way as that of the rest of the sample stratum. The same assumption is made in the subsequent month even though it returns to the sample. Thus no correction is made even if its actual wage growth over the two months turns out to be different from the average of the rest of the stratum. For these reasons matched pairs have not been used in AWE. Instead missing data are imputed and both the responses and the imputed data contribute to the estimate of AWE.

This argument is correct, and ONS is able to give examples of other series which have shown significant drift. On the other hand they have also noted that drift may be less of a problem with a wage or price indicator than it would be with an indicator which measured the volume or value of transactions, such as the retail sales index. This is because market forces might be expected eventually to align wage rates whereas there is no such force present in influencing the sales of one retail outlet compared to another.

The other side of the argument is that, drift not withstanding, the use of matched pairs may produce growth estimates with lower mean square and absolute errors than does the use of the complete sample. ONS recognises that ideally a composite estimator would be used which gave more weight to those firms present in both months but did not ignore the information contained in the responses of the firms present in only one of the months. Indeed ONS staff made a similar observation verbally at the time of the Turnbull King review. Appendix A3.2 contains a note on composite estimation. However ONS also acknowledges that internationally very little use is made of

composite estimation and attributes this to practical difficulties in its implementation. It is difficult to disagree that the method is not useable at present.

If we accept that composite estimation is not possible, then there is the question whether the matched pair approach of the AEI or the imputation approach of AWE is better, given that the aim is to produce a measure of changes to wages over a twelve month or four quarter period rather than an estimate of the level of wages. ONS has not been able to point to any body of experience or work which answers this question clearly one way or the other although there are some attributes of AWE which are clearly more satisfactory- most notably the way in which it handles missing data on bonuses separately from those on ordinary pay . It is possible that the use of matched pairs, which was recommended by Turnbull-King, has led to drift in the AEI estimates. However, the fact that users seem to have been reasonably happy with the series does point to a need for the case for a change to be made clearly on the basis of scientific evidence. Section 5.2 sets out the work which might be done to assess formally the case for a change from the use of matched pairs to the sort of estimation used in the experimental AWE; it also suggests a hybrid in which imputation is used to deal with missing values but matched pairs are used to handle the effects of sample rotation. Other approaches might also be possible such as the use of matched pairs to calculate the growth rate of the average wage for each stratum and use this growth rate to impute a level of the average wage for each missing observation. This imputed value would be used as though it were an actual value in the calculation for the subsequent month, while with the current AEI procedure, each firm is omitted unless it provides data for both adjacent months. It is suggested that the final decision on the appropriate approach should not be made without this work having been completed.

To summarise, the imputation methodology employed by AWE is superior to the matched-pair approach used in AEI in the production of estimates of the level of earnings because it allows for a more representative response and appropriate weighting. But further work is needed to show whether this produces the best estimator of growth in earnings or whether perhaps some combination of the two methods is needed.

Classification

The AWE and AEI are prepared to the same industrial classification. Nevertheless, the same firm may be classified to two different industries in the two different cases. The reason for this is as follows. In the nature of things, as firms change their line of business it becomes appropriate to move them from one industry to another industry. Changes also arise as a result of the ONS receiving new or additional information on the activity of a business that leads to its classification being reassessed, and these are more numerous than those generated by genuine changes in business activity.

The AEI classification is revised each August using the classification shown in the IDBR in the previous January. The logic for making the change in August is that it is a fairly quiet month, being outside of the main ‘bonus season’ months between December and April. The AWE by contrast is revised in January to the January IDBR and large changes in subsequent periods affecting firms employing more than 1000 people are allowed to take effect as they happen. Thus, for the period January to July the AWE classification will differ (and be more up-to-date), for a small proportion of

the sample, from the AEI classification and differences also accumulate from August onwards.

Outlier Detection and Treatment

The Turnbull-King report made a number of suggestions about dealing with outliers in the AEI. These focused on i) the change in the average wage firm by firm from one month to the next and ii) the contribution made by each firm to the movement of the overall index. Noting that there were surprising large movements in average wages reported by some firms from one month to the next, the report recommended (recommendation 17) that all firms whose reported wage changed by a factor of ten or more from one month to the next should be treated as outliers. It also recommended that changes by a factor of three or more should be investigated by the statistician who would decide, in the light of the past responses by the firm and the responses of other firms in the same stratum, whether to classify them as outliers (recommendation 19). Finally, it was recommended that any firm which contributed more than +/- 0.05 per cent per annum to the movement of the overall index should also be treated as an outlier unless the pattern of past returns suggested that this was normal (recommendation 20).

In AWE outlier detection is by way of an automated system which identifies extreme values (both large and small) as those lying more than six standard deviations from the mean of the stratum, with both the mean and the standard deviation evaluated excluding the firm in question. Treatment is by post stratification of outliers into a fully enumerated post-stratum with the appropriate adjustment to register totals to facilitate re-calculation of weights. It should be noted that the outlier system identifies extreme values of the average wage per firm. It is possible, of course, that firms which pay very high or very low wages may show earnings growth which is in line with that for the industry as a whole. This raises the issue as to whether it would be more appropriate to base the outlier detection method on growths rather than levels.

The Turnbull-King Report paid no attention to the question of employment outliers despite the fact that, in both the AEI and AWE, levels of employment affect the calculation of the index. With the AEI they affect the calculation of the average wage in each industry. With AWE they also affect the estimate of employment by the industry and thus the way that the average wages in different industries are combined to give the average wage for the economy as a whole. With AWE in particular, there is therefore a risk that erratic movements in employment in particular firms could lead to shifts in the industry weights which distort the aggregate index.

The ONS is aware of this risk and corrected the omission in the Turnbull-King Report. In the calculation of both indicators, any firm whose employment as reported in MWSS is more than one band from the level reported in the IDBR at the start of the relevant year is treated as an employment outlier. It is given a weight of 1 rather than the stratum weight for the calculation of both the wage bill and the employment total. However, it is not known whether this treatment of employment outliers is adequate. The risk remains that erratic movements in employment arising from outliers are a source of disturbance to AWE but the ONS argue that the chance of this being a problem is small.

Outliers and Revisions

In AEI once a firm is classified as an outlier its status in the month in question is never reviewed. There are, nevertheless, reasons why a review might be sensible. Firms are classified as outliers on the grounds that movements in their wages or, in the case of AWE, the level of their wages is such that they can be assumed to represent only themselves. Judgements about whether firms are actually outliers or not inevitably depend on how extreme their returns appear to be. A firm which is classified as an outlier early on as the returns come in may seem to be less of an outlier when later returns come in and in such circumstances it might be sensible to review the classification.

With AWE this process is automatic. As new data come in both the mean and the standard deviation in the relevant stratum change, so that the outlier status can change. Experience suggests that this can lead to substantial revisions. For example, for December 2006 this process was largely responsible for a revision to AWE of 1.2 percentage points between its first and second estimates. This is discussed further in section 2.6. It should be noted that this problem would be avoided if the outlier/non-outlier distinction were less 'on/off'. Different treatments may well differ in the extent to which they are sensitive to revisions and it would be desirable for this issue to be explored by ONS at some point.

To conclude, outlier identification in AWE is predicated on estimates of levels of pay and employment; an alternative methodology to provide optimal treatment for an indicator of growth should be tested (see section 5.3). Some of the aspects of identification in AWE might be retained, for example, the practice of revisiting outliers at month t+1 and reassessing them in light of subsequent information.

Estimation and Small Businesses

Businesses with fewer than 20 employees, that is small firms, are not sampled in the MWSS, but they are part of the target population for both the AEI and AWE. In the AEI, the average wage for employees in small firms, in each industry, is assumed to be equal to the average wage for employees working in firms in the 20-99 size band. The estimated average wage in each industry is weighted to the target population using fixed employment shares, incorporating the small firms, taken from the IDBR. In AWE the average wage for employees in small firms is assumed to be different to that in large firms. To calculate an estimate of the average wage for small firms, in each industry, the average wage for large firms is calculated then multiplied by an adjustment factor derived from the Annual Survey of Hours and Earnings (ASHE). These adjustment factors are a mean of the ratio of the average wage of employees in small firms to the average wage of employees in large firms. The mean is computed over three years to smooth out fluctuations in the adjustment factors. The estimated average wage for employees in small firms is then multiplied by the employment for those firms taken from the IDBR, to give an estimate of the total wages for small firms. The ASHE covers employees of all businesses and is based on replies pertaining to approximately 160,000 employees annually. The basic estimator for calculating the total earnings and separately the total employment among the larger businesses is a simple separate ratio estimator using employment as the auxiliary variable.

Seasonal Adjustment

Both AWE and the AEI are adjusted seasonally using X12ARIMA. This is done separately for the public and private sector. There is a larger seasonal factor in the estimate which includes bonuses than the estimate of earnings excluding bonuses. The comparison between the two series in this chapter is made using the seasonally adjusted figures because these are used to provide the headline data. However, subsequent analysis of the differences between the two series is carried out using the data without any seasonal adjustment.

Standard Deviations

Following the Turnbull-King Review the ONS did a substantial amount of innovative work to determine the standard deviations of the key outputs from the AEI. At present the theoretical work needed to do the same for AWE has not been completed and ONS has not made a start on the practical implementation of the methods that are needed.

The Statistics Authority needs to take a view on whether it is prepared to authorise AWE as a National Statistic before standard error information is available. The case for waiting for this information is that it is undoubtedly good practice for information on reliability to be provided with data. The case against waiting is that there are many statistics for which no such information is available and AWE would not stand out if approval did not wait for information on standard errors. On the other hand publication of information on standard errors could help build confidence in AWE.

Methodology

The algebraic construction of the two series is described in full in Appendix 2. Here we include worked examples which show how the AEI can differ from the AWE as a result of the failure of the former to take account of changes in the mix of employment. The examples are designed to focus on this aspect of the difference, rather than on the other differences because this issue was the focus of the recommendation in the Turnbull-King Review.

We show three cases. In all cases wage growth is calculated from the average wage in each industry multiplied by employment divided by the average wage in the previous month multiplied by employment. However, with the AEI the employment figures are taken from the IDBR and change only once a year. In this example they are shown as fixed. By contrast, with AWE the employment figures are those calculated from the MWSS itself. In the first example there is no change in the pattern of employment, either from the IDBR to the first month of the comparison or from the first month to the second month. This means that the growth rates shown by the AEI and AWE are the same.

In the second example the average wage does not change in either industry, with the consequence that the AEI is unchanged. However, employment in the low-paying industry rises and that in the high-paying industry falls. As a result the average wage in the economy declines. This is reflected in the AWE measure.

The third example combines the previous two. Wage rates change and there is a shift in employment between the first and second months. The AEI shows the same growth

as in Case 1. The AWE shows lower growth because the calculation reflects not only the increase in wages but also the shift in employment from industry two to industry one. It should be noted that the growth in AWE in the third example is approximately equal to the growth in the AEI plus the growth in AWE calculated in Case 2 where the wages remain unchanged and the impact arises solely through employment composition effects.

Overall, the AWE method is to be preferred since it gives a better estimate of the movement of the 'true' average. On these grounds AWE is logically superior to AEI. This is the driver behind Finding 1 presented subsequently.

Table 2-1 An Illustrative Calculation of the AEI and AWE showing the Effect of Changes in the Employment Mix.

Case 1		**Weighting Effects**	
		Industry 1	Industry 2
Employment IDBR		10000	30000
Month 1	Wage	400	500
	Employment	10000	30000
Month 2	Wage	420	510
	Employment	10000	30000
AEI Wage Growth		2.632%	
AWE Wage Growth		2.632%	

Case 2		**Composition Effects**	
		Industry 1	Industry 2
Employment IDBR		10000	30000
Month 1	Wage	400	500
	Employment	10000	30000
Month 2	Wage	400	500
	Employment	11000	29000
AEI Wage Growth		0.000%	
AWE Wage Growth		-0.526%	

Case 3		**Combined Effects**	
		Industry 1	Industry 2
Employment IDBR		10000	30000
Month 1	Wage	400	500
	Employment	10000	30000
Month 2	Wage	420	510
	Employment	11000	29000
AEI Wage Growth		2.632%	
AWE Wage Growth		2.158%	

The distinction between AWE and AEI can also be shown algebraically. Here

$w_{hk,t}$ is the average wage computed in industry *hk* in month *t*
$\theta_{hk,t}$ is the share of industry *hk* in total employment in month *t*

h is the industrial classification for a stratum, defined by 2-digit SIC combined with public/private classification. There are 116 public/private industries.

k represents the components of the 20-industry classification at which the AEI is published. The distinction between broad and narrow classifications is not relevant in this section but is required at a later state.

The average wage in the AEI is computed as
$$w_t^{AEI} = \sum_{h,k} w_{hk,t}\theta_{hk,T}$$
where T indicates the base period for which the employment weights are calculated.

The average wage in AWE is calculated as
$$w_t^{AWE} = \sum_{h,k} w_{hk,t}\theta_{hk,t}$$
where θ_{lt} is the employment share as calculated from MWSS in month *t.*

The relevant growth rates are, with $\Delta w_{hk,t} = w_{hk,t} - w_{hk,t-1}$
$$\frac{\Delta w_t^{AEI}}{w_{t-1}^{AEI}} = \frac{\sum_{hk} \Delta w_{hk,t}\theta_{hk,T}}{\sum_{hk} w_{hk,t-1}\theta_{hk,T}} \qquad (2.1)$$
and
$$\frac{\Delta w_t^{AWE}}{w_{t-1}^{AWE}} = \frac{\sum_{hk} \Delta w_{hk,t}\theta_{hk,t-1}}{\sum_{i} w_{hk,t-1}\theta_{hk,t-1}} + \frac{\sum_{i} w_{hk,t-1}\Delta\theta_{hk,t}}{\sum_{i} w_{hk,t-1}\theta_{hk,t-1}} + \text{ second order terms} \qquad (2.2)$$
Since the shares always add to one, it follows that
$$\sum_{h,k} \Delta\theta_{hkt} = 0$$
so that there is no composition effect if the average wage in period *t* is the same in all industries.

In Case 1 of table 2.1 $\theta_{hk,t-1}=\theta_{hkt}=\theta_{hkT}$ so that the growth rate of AWE is the same as that of AEI. In the second case $\theta_{hkt}\neq\theta_{hk,t-1}=\theta_{hkT}$ but $w_{hkt}=w_{hk,t-1}$ so that changes to the AWE arise only from changes in the employment mix, and there is no change to the AEI. In the third case both the change in the wages in each industry from the first case and the change in employment from the second case are present. If the overall change from month 1 to month 2 involves both these changes, then the wage growth of AWE shown in case 1 represents the first term on the right hand side of equation (2.2) and case 2 represents the second term. There is a residual of 2.158-(2.632-0.526)=0.052 percentage points. The example shows how AWE indicates the change in the average wage taking account of employment mix as well as the effect of wage changes in each industry. The AEI omits the employment mix term.

2.4 Finding 1: The Relative Merits of AWE and AEI

The fact that, as demonstrated above, AWE provides a measure of average pay in the economy while AEI does not leads to Finding 1.

The AWE offers in principle a better measure than the AEI of changes in the average wage in the economy. However, subsequent findings indicate that more work is desirable to ensure that it is also in practice a better measure.

2.5 The Relationship between the Two Series

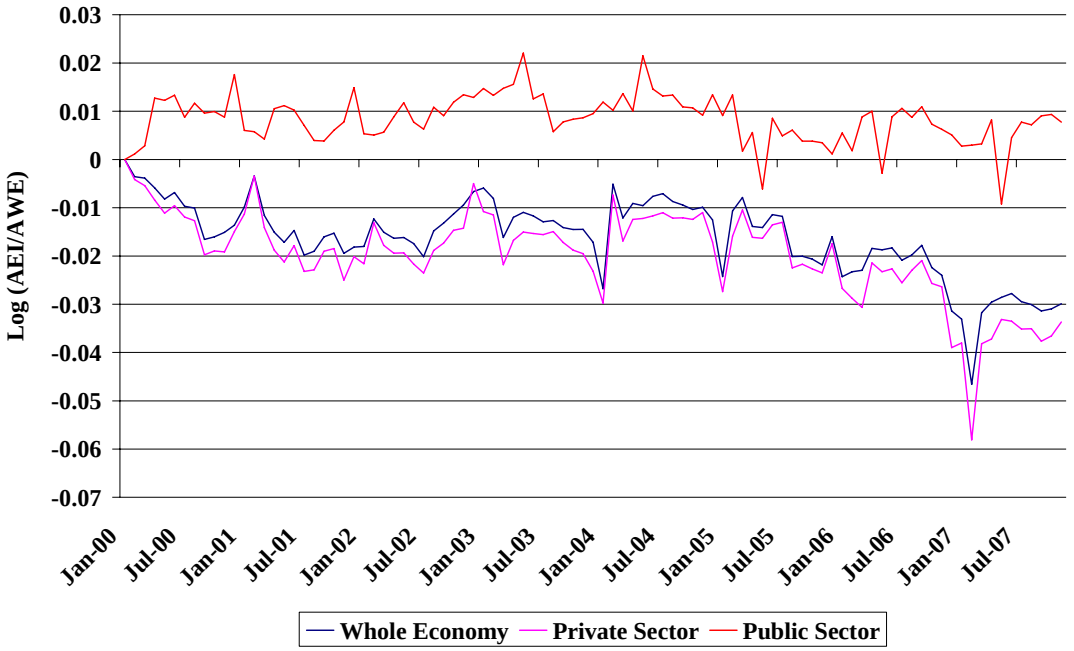
When the data are released users are encouraged to focus on the rate of change between the average of each indicator in the most recent three months available and the analogous figure a year earlier. This indicator has the advantage that it smooths out measurement errors and other short-term disturbances and, for users who do not wish to do any smoothing for themselves, is a useful summary indicator. Divergences between the growth rates for the two series, thus defined, have been a source of concern to users and contributed to the need for this review.

Nevertheless, such a comparison is not very helpful if one wishes to understand the statistical nature of the relationship between the two series. Suppose, for example, that the series have the property that shocks tend to persist[1] and that there is a blip in one month's data. The fact that the headline figure is calculated by comparing a three-month average with the average for the same period a year earlier means that the blip takes three months to build up to its full effect. It persists at its full effect for ten months in total before fading away over the next two months, so that it influences the headline rate for fourteen months in total. If different methods of calculation lead to blips occurring in one series rather than the other, or in the two series at different times, then a comparison of changes over twelve months creates an impression of persistent disturbances. This persistence arises not because the same factor is present in all twelve months of the data but because attention is focused on a change over twelve months. Thus it is more helpful to focus either on monthly growth rates of the series or on cumulated changes over time i.e. on the levels of the series. Looking at the logarithms of the series does this in a way which is not affected by the scale of the data

A first impression can be gained by looking to see how the indices have moved since AWE was first calculated in 2000. Figure 2-1 shows the log of the AEI relative to AWE with the difference adjusted to zero in January 2000.

[1] In statistical terminology the series of earnings should be expected to be I(1). This means that the change in the series is a random process, probably including a trend term reflecting the way in which earnings tend to rise over time. A sharp rise in one month which is not offset in the next month leads to a high annual rise, but is not, in this case followed by a below annual rise in the next year. This view of the data is consistent with the way in which inflation targeting affects the price level; if inflation is above target in one year, the Bank of England does not attempt to deliver inflation below target in the next year. Instead prices are permanently higher.

Figure 2-1 AEI and AWE Compared



Data: AEI: Series LMNQ LNKY LNNJ SA Including Bonus
AWE: AEI Supplementary Tables2a SA Inc Bonus Ex Arrears.

The general pattern is of the two indicators gradually drifting apart; the negative value of the difference for the private sector and the whole economy indicates that for these, average weekly earnings have risen more quickly than the average earnings index, with the cumulative gap between January 2000 and November 2007 amounting to 2.9 per cent for the economy as a whole and 3.3 per cent for the private sector. The drift is just under 0.4 per cent per annum for the economy as a whole and just over this for the private sector. The two public sector measures of earnings show much less sign of drift, with an average drift over the period of under 0.1 per cent per annum but with the average earnings index rising faster than average weekly earnings.
If someone moves from a low-paying firm to a high-paying firm the measure of average weekly earnings should record an increase in average earnings, while the average earnings index records no change. Thus the discrepancy observed for the whole economy and the private sector implies a drift from low-paying to high-paying firms. However, there are a number of other factors behind the difference and it is wrong to conclude that such a movement is the main influence on or explanation of the divergence between the AEI and AWE.

An analysis of the relationship between the monthly growth rates of the series and their relationship with growth rates drawn from other indicators follows in Chapter 3.

2.6 Appendix to Chapter 2. The Interaction between Outliers and Revisions

During the processing of the data for January 2007 the AWE for December 2006 was revised up from £418.0 to £422.7 as a result of late returns. This was an increase in the year on year growth of 1.2 percentage points.

Although the employment estimates underlying AWE were revised the main reason for the revision to AWE was the revision to the total wages and salaries estimate – this was revised upwards by over £140m or 1.3 percentage points.

Of the £140m increase, nearly £90m came from one stratum – the lowest stratum (20 to 99 employment) in Div 67. This was because the late arrival of data from one firm meant that another very high-paying firm which had previously been classed as an outlier was no longer regarded as such.

Table 2-2 shows how this can happen, using illustrative data to avoid the risk of disclosure. Working from the provisional set of data firm 21 is classified as an outlier because its pay per employee is more than six standard deviations from the mean for the rest of the data set (with both the mean and the standard deviation calculated excluding the firm). Once a reply is received from firm 3 both the mean and the standard deviation are increased. Firm 21 is no longer more than six standard deviations from the mean. As a result it is no longer classified as an outlier and its contribution to the average earnings figure for the whole economy is therefore increased many times. These effects are not present in the AEI because judgments about outliers are not revised on receipt of new data. There is a good case that they should be although at the same time a method of treating outliers which results in revisions of this type is not fully satisfactory.

Table 2-2 An Illustration of the Effect of Revisions on Outliers

	Pay per Employee (£ per week)	
Firm	Provisional	Revised
1	272	272
2	423	423
3	Not available	2,841
4	355	2,073
5	156	156
6	176	176
7	219	219
8	91	91
9	941	941
10	98	98
11	141	141
12	236	236
13	132	132
14	268	268
15	241	241
16	155	155
17	203	203
18	930	930
19	4,786	4,786
20	890	890
21	6,627	6,627
22	130	130
23	464	464
24	198	198
25	175	175
Mean	508	677
St dev	968	1,099
Mean+6 St dev	6,570	7,510

Chapter 3. Comparison with other Data

3.1 Introduction

There are a number of indicators of movements in earnings in addition to the AEI and AWE. In this chapter we explore the relationship between these other measures of earnings and the two indicators of particular interest so as to establish whether the relationship with any of these other indicators might influence a choice between the AEI and AWE.

The analysis is conducted in a manner which reflects the data to which the AEI and AWE are being compared. The correlations between AEI and AWE are also shown calculated on the same basis. They differ from example to example because on some occasions seasonally adjusted data are used while on others the correlations are calculated after using regression to remove seasonal effects. Some variables are quarterly and others are monthly and the relevant time periods differ. All of these factors influence the correlations between AWE and AEI.

3.2 The Annual Survey of Hours and Employment

The Annual Survey of Hours and Earnings (ASHE) is compiled from a 0.8% sample[2] of national insurance numbers extracted from the PAYE system. As a result it excludes the majority of people who are below the national insurance threshold. The survey produces, among other things, an estimate of the average weekly wage that can be compared directly with AWE. The growth in both variables can be compared with the AEI index. The relevant data are shown in Table 3-1. All data include bonuses and arrears and are not seasonally adjusted. Given the small number of observations correlations are not shown.

Table 3-1 Growth Rates of Average Earnings Measured by ASHE, AWE and AEI

	ASHE	AWE	AEI	ASHE	AWE	AEI
2000	354.5	316.2	98.5	Annual Growth		
2001	373.8	334	103.3	5.4%	5.6%	4.9%
2002	391.5	346.1	107.2	4.7%	3.6%	3.8%
2003	403.9	354.7	110	3.2%	2.5%	2.6%
2004	413.6	368.7	115	2.4%	3.9%	4.5%
2005	422.8	384.4	119.9	2.2%	4.3%	4.3%
2006	441.8	399.9	124.1	4.5%	4.0%	3.5%
2007	452	417.1	128	2.3%	4.3%	3.1%
RMSE					1.33%	1.31%

Data ASHE:
2000-2003 http://www.statistics.gov.uk/articles/nojournal/ASHE_analysis.pdf. Table 1 with imputation
2004-2007 Table 1.1a
2004 http://www.statistics.gov.uk/downloads/theme_labour/ASHE_2004_inc/2004_all_employees.pdf
2005 http://www.statistics.gov.uk/downloads/theme_labour/ASHE_2005/2005_all_employees.pdf
2006 http://www.statistics.gov.uk/downloads/theme_labour/ASHE_2006/2006_all_employees.pdf
2007 http://www.statistics.gov.uk/downloads/theme_labour/ASHE_2007/2007_all_employees.pdf
AWE: AWE_Tables_Updated.xls. April Figures Whole Economy
AEI: Series LNMM April Figures

[2] Until 2006 the sample was 1% of national insurance numbers.

The wage levels shown in ASHE are very appreciably higher than those shown in AWE; this may well be because, as noted above, people whose earnings are below the national insurance threshold (e.g. because they work part time) are not included in ASHE while they are present in AWE.

The growth rates of the three series are shown. We summarise the differences between these by looking at the root mean square deviation (RMSE) of AWE and the AEI from ASHE. Plainly there is little to choose between the two indices. ASHE does not provide a basis for saying that one of them is to be preferred.

3.3 PAYE Returns

Her Majesty's Revenue and Customs (HMRC) has made available monthly pay as you earn (PAYE) returns of income tax and national insurance contributions. These obviously cover a different variable, total tax paid on wages and salaries rather than the average wage. As well as the fact that the series is affected by changes in total employment, it includes tax on redundancy pay and on arrears. Nevertheless some comparison can be made.

HMRC advises that the PAYE data reflect earnings in the previous month; thus the May return corresponds to April earnings. Employers do, however, have some flexibility in when they make their returns so the lag is not exact. A further problem in making the comparison is that the HMRC data are not seasonally adjusted. Indeed the series of "cleaned up" data runs only from May 2005 to January 2008; normally three years' data are needed before the standard seasonal adjustment methods can be used. The comparison made here uses a regression method to take account of the seasonal effects, but this, too may be affected by the very short length of the series. We regress separately the first differences of the log values of the PAYE series, the AEI and AWE on monthly dummies and use the residuals of the resulting regressions to represent the non-seasonal components of the three series. The resulting correlations are shown in Table 3-2.

Table 3-2 Seasonally Corrected Correlations between Δlog PAYE Total Revenue, Δlog AEI and Δlog AWE: Whole Economy

	AEI	AWE
PAYE	0.32	0.51
AEI		0.82

Data: PAYE. Supplied by HMRC. May 2005 to Jan 2008 (related to April 2005-December 2007).
AEI. LNMM April 2005-December 2007
AWE AWE_Tables_Updated.xls Whole Economy April 2005-December 2007

Neither correlation with the PAYE series is high. Nevertheless, for the period explored, AWE matches the PAYE data better than does the AEI. It has to be remembered, however, that the PAYE figures reflect employment changes as well as changes in average wages. There are no monthly data which allow us to examine this.

HMRC also provided data for PAYE receipts for financial intermediation, together with the AEI and AWE data for the same sector. The correlations[3] are shown in Table 3-3.

[3] AWE data to the relevant definition were available only to November 2007. Thus the seasonal adjustments and then the correlations are calculated on the monthly changes from May 2005 to November 2007.

Table 3-3 Seasonally Corrected Correlations between Δlog PAYE Total Revenue, Δlog AEI and Δlog AWE Measures of Monthly Growth in Average Earnings: Financial Intermediation

	AEI	AWE
PAYE	0.15	0.63
AEI		0.74

Data: PAYE. Supplied by HMRC. May 2005 to Jan 2008 (related to April 2005-December 2007).
AEI. JVUT April 2005-December 2007
AWE Data in HMRC Spreadsheet. April 2005-December 2007

Here too AWE matches the PAYE series better than does the AEI; indeed it does it slightly better than for the economy as a whole.

3.4 Hours Worked

The Labour Force Survey (LFS) provides quarterly estimates of hours worked. Other things being equal, an increase in hours worked per person employed should be associated with an increase in Average Weekly Earnings. It need not lead to an increase in the Average Earnings Index if it is caused by people moving from low hours employers to high hours employers. But such shifts are unlikely to be a major cause of the fluctuations in hours worked; examination of the relationship between changes to hours worked and changes to quarterly averages of the two earnings measures might therefore offer an indication of whether one was better than the other.

As with the analysis of PAYE data we look at correlation coefficients between changes in the log values of the two earnings measures with changes in the log of hours worked after using regression to remove seasonal effects. In both cases the correlations are very weak.

Table 3-4 Seasonally Corrected Correlations between Δlog Hours Worked and Δlog AEI and Δlog AWE. 2000Q2-2007Q4.

	AEI	AWE
Hours Worked	0.049	0.058

Data: Hours Worked EAADM
AWE: AWE_Tables_Updated.xls AWE Sect Monthly
AEI: LNMM

3.5 Labour Force Survey

The Labour Force Survey also provides quarterly estimates of average weekly earnings which can be compared with the AWE and AEI. Until 2004 LFS data are available only biannually. Table 3-5 shows the differences between these measures of mean weekly earnings. In this table the AEI figures are converted from an index into £ per week by assuming that, in the first quarter of 2000, they match AWE exactly.

The general impression from these data is that, for the most part, the numbers are very similar. The Labour Force Survey does not show the first quarter peaks in earnings which reflect bonuses. There are a number of reasons for this. First, bonuses are not mentioned explicitly when collecting the data, with the likely consequence that people leave them out and secondly response among people with large bonuses may be low. Nevertheless from 2006Q4 there is also an impression that the AWE tends to drift

away from the LFS/AEI measure. This is a surprise since the AEI might be thought to be more susceptible to drift.

Table 3-5. Average Weekly Earnings from the Labour Force Survey, the AWE and the AEI

	LFS	AWE	AEI
2000 Q1		325	325
Q2		316	317
Q3		318	317
Q4		324	325
2001 Q1		344	342
Q2	331	333	332
Q3		333	331
Q4	340	337	336
2002 Q1		356	352
Q2	344	346	344
Q3		344	342
Q4	347	347	349
2003 Q1		367	364
Q2	353	355	354
Q3		355	355
Q4	359	360	360
2004 Q1		386	383
Q2	370	370	370
Q3		369	369
Q4	372	374	376
2005 Q1	377	405	402
Q2	382	385	385
Q3	386	386	383
Q4	388	391	389
2006 Q1	395	427	419
Q2	394	404	401
Q3	401	402	398
Q4	402	409	404
2007 Q1	410	454	438
Q2	412	421	414
Q3	413	422	414
Q4	416	427	420

Data: LFS. Supplied by ONS
AWE: AWE-Tables_Updated. AWE Sect Quart. Whole Economy
AEI LNMM. 2000Q1 scaled to AWE value in 2000Q1

It is in fact difficult to know how much importance to give to this, since if we look at the changes over the two-quarter periods for which the time-series goes back further (Q2/Q4 and Q4/Q2) in the logs of the three measures, then the RMSE of the deviation of AWE from the LFS measure is 0.834% while for the AEI measure it is 0.840%. When we look at the quarterly changes from 2005Q1 onwards and use seasonal dummies to remove seasonal effects, we find that the correlations are very low. This is shown in Table 3-6.

The ONS report that the standard error in the LFS figures for 2007 Q4 is £2.51; the AEI figure lies close to two standard deviations from the LFS figure while AWE figure lies over four standard deviations from this. Nevertheless, there must be some questions about how well the LFS picks up bonus payments and it is likely to underestimate mean pay in the economy.

Table 3-6 Seasonally Corrected Correlations between Δlog LFS wage and Δlog AEI and Δlog AWE. 2005Q2-2007Q4.

	AEI	AWE
LFS	0.034	0.153
AEI		0.954

Data: LFS. Supplied by ONS
AWE: AWE-Tables_Updated. AWE Sect Quart. Whole Economy
AEI LNMM.

Thus the short-term movements in the LFS series do not match either of the other earnings indicators. By contrast, once the seasonal effects were removed, the correlation between the quarterly growth rates in the AEI and AWE was 0.954.

3.6 Report on Jobs

There are a number of other series with which AEI and AWE may be compared. The *Report on Jobs* produced by NTC includes a monthly survey which asks respondents "Are average salaries awarded to staff placed in permanent positions higher, the same or lower than one month ago". The report publishes an indicator calculated as 50 plus percentage reporting higher minus percentage reporting lower and interprets a figure of 50 or above as indicating the rate of inflation.

The time series of these data has been compared with the changes in the AEI and AWE for the private sector excluding bonuses. These figures are inevitably much less volatile than the figures including bonuses. Bonuses have the effect of movements in both series which are large compared with other changes; when these are taken out the impact of methodological differences becomes relatively larger and this depresses the correlations compared with what would be found with series which included bonuses[4].

Table 3-7: Correlations between the Growth in the Two Measures of Earnings and the Data produced by the Report on Jobs. February 2000-December 2007. Seasonally Adjusted Data Used.

	AEI	AWE
Report on Jobs	0.07	0.11
AEI		0.55

Data: Report on Jobs Data Seasonally Adjusted: REC Permanent Salaries
AEI: JQEC
AWE: AWE_Supplementary _tables.xls Table 2b - SA ex bon, ex arr by sect. Pvt Sector

[4] If the same or two similar highly variable series is added to two series which are relatively poorly correlated, the correlation between the two series is normally substantially increased. It is for this reason that the series including bonuses are more closely correlated than those excluding bonuses.

Little can be learned from this, but one would not expect the indicator to be closely related to pay pressures in the economy as a whole.

3.7 Settlements Data

There are a number of other series with which comparison is possible but this is complicated by the fact that the series in question are published as changes over twelve months or, as with the headline AEI, as the average of the last three months compared to the average of the same period one year earlier. As we noted above, an analysis of the relationship between such series, whether visual or using correlation/regression methods, tends to produce a misleading picture. One method of avoiding this problem is to look at correlations between the variables over intervals that do not overlap. Thus if a series shows changes over twelve months, the December to December growth rates can be compared. This means that, instead of having the ninety-five monthly changes available with genuinely monthly data, there are only six available to make the comparison. As a result little can safely be inferred.

Where figures are provided for changes over twelve months (rather than three month averages relative to earlier three-month averages) a practical alternative suggests itself, although we have not examined its formal statistical properties. For 2000 we take the average of the logarithms of the two earnings measures and work forward from this using the twelve-month growth rates. We use regression methods as before to remove seasonal effects from the series and then compute the correlations. This method allows us to compare the settlements data with both earnings series. The resulting correlations are shown in Table 3-8. These provide no reason for choosing one measure or the other and suggest, perhaps not very surprisingly, that short-term movements in earnings are not closely related to short-term movement in settlements. The correlation between the two measures of earnings on the same basis is 0.927.

Table 3-8 Correlations between Δlog Settlements and Δlog AEI and Δlog AWE. Feb 2000-Dec 2007. Seasonal Effects Removed by Regression. Pattern for settlements data for 2000 inferred from mean of log AWE and log AE in 2000

	AEI	AWE
Settlements	0.162	0.157
AEI		0.927

Data: Settlements Data: Bank of England
AEI: LNMM
AWE: AWE: AWE-Tables_Updated. AWE Sect Monthly. Whole Economy

3.8 The National Accounts

The AEI provides the data for the construction of the measure of wages and salaries paid to households ahead of the availability of data constructed from tax records. Once the data are drawn from tax records, there are some differences of definition; for example share awards and share incentive plans are a part of wages and salaries for national accounting purposes but are not counted as earnings. At present the tax records are used to compile the wages and salaries variable only up to the end of 2004. The analysis is therefore carried out only up to 2004Q4.

Table 3-9 Correlations between Δlog Wages and Salaries from the National Accounts and Δlog AEI and Δlog AWE. 2000Q2 to 2004Q4. Seasonally Adjusted Data

	AEI	AWE
National Accounts	0.525	0.496
AEI		0.654

Data: Wages and Salaries: ROYJ
AEI; LMNQ
AWE: AWE Tables Updated AWE Sect Quarterly: Whole Economy

Table 3-9 points to a very slightly higher correlation for the AEI than for the AWE but, as with the earlier comparisons it is not possible to come to any firm conclusions. Of course the national accounts figures reflect movements in employment as well as movements in wages. To explore the implications of this we deduct from the change in log wages and salaries the change in the log of employee jobs (series BCAJ). We then find that the correlation for AEI falls to 0.518 while that for AWE falls further, to 0.416. This is not definitive but it is a source of some concern and is consistent with the idea that the current AWE is designed primarily to provide a good measure of the level of wages (see section 5.6) rather than a good measure of the short-term growth of wages.

3.9 Finding 2: Alternative Indicators of Wage Change

There is a range of indicators of wage movements available as alternatives to AWE and AEI. Some of the indicators explored are slightly more closely related to the AEI than AWE while for other indicators the opposite is true. There is some concern that the AEI tracks the movement of wages and salaries per employee in the national accounts better than does AWE. But the most important point to observe is that, when looking at monthly or quarterly growth in earnings, the AEI and AWE are much more closely related to each other than they are to any of the other measures. Thus a general conclusion which follows from this section is that other data series do not provide any solid guidance as to whether AWE should be preferred to the AEI or vice versa.

Chapter 4. Causes of the Divergences between the Series

4.1 Introduction

This chapter shows cumulatively how one can progressively change AWE into the AEI by putting the treatment of various issues on a common basis and making other changes such as removing employment composition effects. The chapter builds heavily on the work presented in Parkin (2008) but examines in more detail the effects of the composition of employment so as to be able to link what is observed to the decomposition presented in equation (2.2).

In the analysis of Chapter 2 we looked at the two measures seasonally adjusted because they provide the basis for the headline figures. However, to understand the influences on the difference between them it is important to work with the raw, unadjusted data. We do that in this chapter.

4.2 Different Measures of Earnings Growth

Eight variants of the earnings index are considered

1. AWE including bonuses and arrears not seasonally adjusted.
2. AWEWage - the wage-change component of AWE after the employment effects and the residual term are removed (see page 15)
3. AWES- the wage-change component after the imputation for small firms is removed
4. AWESM- AWES recalculated using the matched-pair approach of the AEI.
5. AWEM- AWESM with the responding firms reclassified on an AEI basis.
6. AWEIM- AWEM with outliers treated in the same way as in the AEI rather than using the AWE treatment, It should be noted that the treatment of outliers is frozen in variants 1 to 5
7. AEIW- the AEI calculated using AWE employment weights applied at the twenty-industry level.
8. AEI itself.

The logic behind this is as follows. If there were no methodological differences between AWE and AEI except that arising from the fact that AWE uses actual employment rather than fixed values, then the difference between AWE and AWEWage would represent the impact of changes in the composition of employment together with the residual of equation (2.2) and the difference between AWEWage and AEI would represent the impact of using employment weights drawn from the MWSS instead of fixed values.

However, as Chapter 2 has made clear, there are also differences arising for a number of other reasons. The move from AWEWage to AWEIM undoes these differences in a sequence which is in one sense arbitrary, but in another sense is dictated by the differences between the approaches. Of the various differences, the sequence is, at least to some extent dictated by the nature of the changes. For example it is not possible to use the AEI classification of outliers in a dataset which includes firms not present in the AEI, and therefore the matched pair approach has to be introduced before the AEI classification of outliers is used. AEIW is an intermediate state

between AWEIM, where the AEI industry wage rates are aggregated using AWE employment figures and the AEI, where they are aggregated using fixed weights. In AEIW they are aggregated to the 20-industry level using AEI weights. AWE weights are then used to produce the aggregate. This gives an indication of the importance of differences in weights at a broad level compared to the effects of differences in weights at a narrow level.

The subscript *h* is used to indicate industries at the 116 industry level and *k* at the 20-industry level. Thus $w_{hk,t}$ is the average wage in industry *h* of broad category *k* in month *t*. The monthly growth rates used in the calculations are derived as follows.

The average wage in AWE is calculated as $w_t^{AWE} = \sum_{h,k} w_{hk,t}^{AWE} \theta_{hk,t}$ where $\theta_{hk,t}$ is the employment share as calculated from MWSS in month *t*. $w_{hk,t}^{AWE}$ is the average wage in industry *h* of broad category *k* in month *t* calculated on an AWE basis.

The growth rate of AWE is

$$\frac{\Delta w_t^{AWE}}{w_{t-1}^{AWE}} = \frac{\sum_{h,k} w_{hk,t}^{AWE} \theta_{hk,t}}{\sum_{h,k} w_{hk,t-1}^{AWE} \theta_{hk,t-1}} - 1$$

The fact that matched pairs are used in the calculation of AEI means that there are two values of the average wage for each month, depending on whether it is being compared with the following or the preceding month. $w_{hk,t-1,t}^{AEI}$ is used to indicate the wage in period *t-1* calculated from the matched sample also used in period *t*. We use a similar notation for the other variants which are computed using matched pairs and where the value of the average wage in any period depends on whether it is being compared to the previous or the following month. In variants where the employment weights are re-evaluated in each month they too may depend on whether the comparison is with the previous or the following month and this is indicated appropriately. With the AEI this is not an issue since the employment weights are fixed to those of month *T*, with only intermittent rebasing.

The growth rate of AEI is

$$\frac{\Delta w_t^{AEI}}{w_{t-1}^{AEI}} = \frac{\sum_{h,k} w_{hk,t}^{AEI} \theta_{hk,T}}{\sum_{h,k} w_{hk,t-1,t}^{AEI} \theta_{hk,T}} - 1$$

Working from the linear decomposition of AWE growth (equation 2.2) we define AWEWage as

$$\frac{\Delta w_t^{AWEWage}}{w_{t-1}^{AWEWage}} = \frac{\sum_{h,k} \Delta w_{hk,t}^{AWE} \theta_{hk,t-1}^{AWE}}{\sum_{h,k} w_{hk,t-1}^{AWE} \theta_{hk,t-1}^{AWE}} \qquad (4.1)$$

The other variants of AWE are calculated by replacing $w_{hk,t}^{AWE}$ and $\theta_{hk,t}^{AWE}$ by alternative measures of the wage in each industry, and, as with AWEWage, focusing only on the wage component. Thus

$$\frac{\Delta w_t^{AWES}}{w_{t-1}^{AWES}} = \frac{\sum_{h,k} \Delta w_{hk,t}^{AWES} \theta_{hk,t-1}^{AWES}}{\sum_{h,k} w_{hk,t-1}^{AWES} \theta_{hk,t-1}^{AWES}} \tag{4.2}$$

$$\frac{\Delta w_t^{AWESM}}{w_{t-1}^{AWESM}} = \frac{\sum_{h,k} \Delta w_{hk,t}^{AWESM} \theta_{hk,t-1}^{AWESM}}{\sum_{h,k} w_{hk,t-1}^{AWESM} \theta_{hk,t-1}^{AWESM}} \tag{4.3}$$

$$\frac{\Delta w_t^{AWEM}}{w_{t-1,t}^{AWEM}} = \frac{\sum_{h,k} \Delta w_{hk.t}^{AWEM} \theta_{hk,t-1,t}^{AWEM}}{\sum_{h,k} w_{hk,t-1,t}^{AWEM} \theta_{hk,t-1,t}^{AWEM}} \tag{4.4}$$

$$\frac{\Delta w_t^{AWEIM}}{w_{t-1,t}^{AWEIM}} = \frac{\sum_{h,k} \Delta w_{hk,t}^{AWEIM} \theta_{hk,t-1,t}^{AWEIM}}{\sum_{h,k} w_{hk,t-1,t}^{AWEIM} \theta_{hk,t-1,t}^{AWEIM}} \tag{4.5}$$

There remain differences between $w_{hk,t}^{AWEIM}$ and $w_{hk,t}^{AEI}$. These arise because the size band estimates in AWEIM (and AWE) are combined using estimates of employment in each size band complied directly from the IDBR, while in AEI the size-band estimates are combined using base-period IDBR data updated using the MWSS. Thus for practical purposes the comparison between AWEIM, AEIW and AEI indicates the differences between aggregating using employment weights which are always up to date, employment weights which are fixed at the narrow level but up to date at the broad level and weights which are derived from base-period IDBR data. We have

$$\frac{\Delta w_t^{AEIW}}{w_{t-1,t}^{AEIW}} = \frac{\sum_k \theta_{kT}^{AWE} \left(\frac{\sum_h \Delta w_{hk,t}^{AEI} \theta_{hk,t-1,t}^{AEI}}{\sum_h \theta_{hk,t-1,t}^{AEI}} \right)}{\sum_k \theta_{kT}^{AWE} \left(\frac{\sum_h w_{hk,t-1,t}^{AEI} \theta_{hk,t-1,t}^{AEI}}{\sum_h \theta_{hk,t-1,t}^{AEI}} \right)} \tag{4.6}$$

where

$$\theta_{kt}^{AWE} = \sum_h \theta_{hkt}^{AWE}$$

In the first part of this chapter we present the key summary statistics for the different series based on month on month changes. We then show graphically the various series. Here we apply rolling annual growth rates computed from quarterly averages, so as to put them on the same basis as the standard headline data.

Table 4-1 Summary Statistics for the Different Measures of Wage Growth February 2001-January 2008

	AWE	AWEWage	AWES	AWESM	AWEM	AWEIM	AEIW	AEI
Average	0.390	0.395	0.396	0.400	0.424	0.381	0.378	0.376
StDev	2.666	2.643	2.612	2.689	2.752	2.534	2.502	2.429
Correlation with AWE		0.998	0.997	0.993	0.987	0.978	0.975	0.973

Table 4-1 shows the summary statistics for the eight measures. The series are all strongly correlated with AWE (and therefore also with each other). It is also worth noting that the evidence for drift arising from the use of matched pairs is only limited. AWESM, computed using matched pairs, grows at an average rate of 0.004 percentage points per month faster than AWES which does not use matched pairs. The other changes have much larger influences on the estimated growth rate of wages. The most marked point in the table is that the average growth of AWEIM is appreciably (but not statistically significantly) below that of AWEM. This could be an indication that the AEI treatment of outliers- used in AWEIM but not AWEM- introduces a bias to the growth rate.

The correlations in Table 4-1 are exaggerated by the presence of seasonal effects. Table 4-2 shows these after seasonal terms have been removed.

Table 4-2 Summary Statistics with Monthly Patterns Removed. February 2001- January 2008

	AWE	AWEWage	AWES	AWESM	AWEM	AWEIM	AEIW	AEI
St dev	1.033	1.025	0.988	1.037	1.110	0.920	0.966	0.922
Correlations		0.996	0.996	0.971	0.927	0.903	0.900	0.909

The mean is now taken up in the seasonal dummies. The differences in the standard deviations are not large and of no obvious importance- although the claim that the use of matched pairs generally results in smoother data than the procedure adopted in the AWE is not borne out. The correlations are, not surprisingly, lower than those shown in the raw data. But nevertheless, they remain high.

4.3 The Time-series Patterns of the Different Estimates

The analysis so far has been in terms of monthly growth rates, and the reasons for looking at the data in this way rather than at the headline presentation were discussed in section 2.5. However, in order for users to look at the consequences of the different adjustments it is much more helpful to look at the headline growth rates, as annual changes to rolling quarterly averages.

Figure 4-1 AWE, AEI and AWE wage component. Annual Changes to Rolling Quarterly Averages.

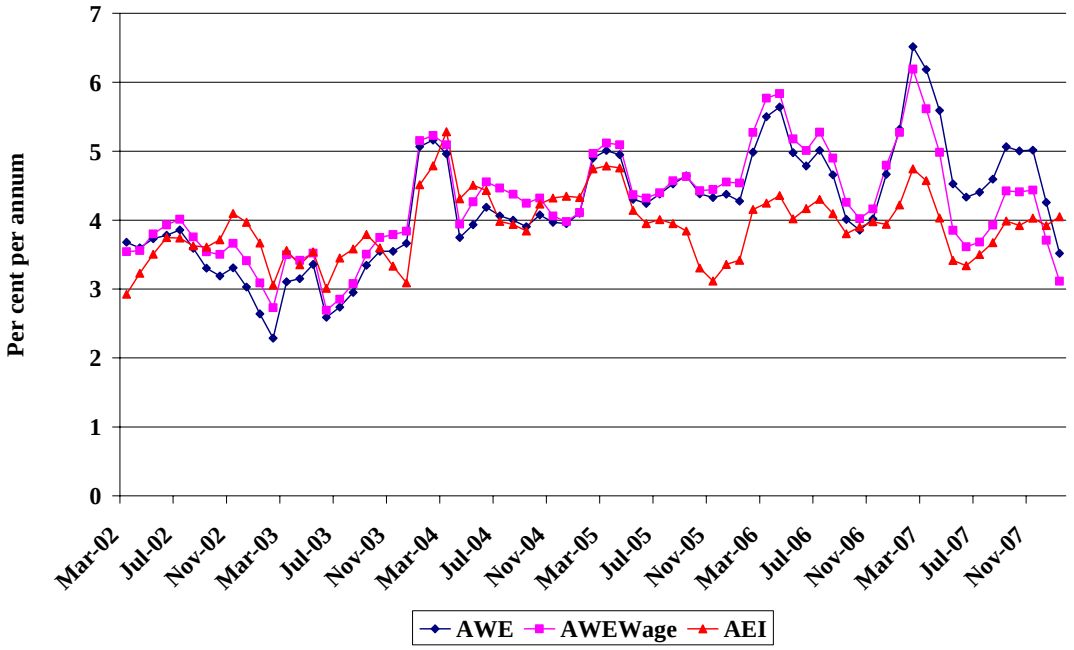
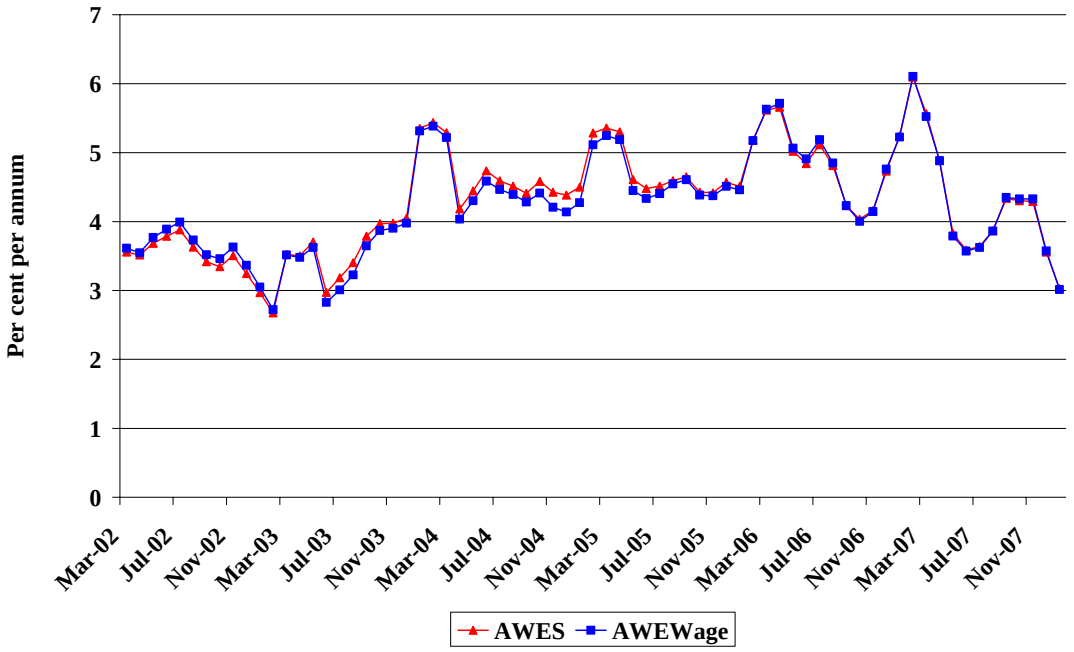


Figure 4-1 shows the "headline" figures for the growth rate between rolling quarterly averages and those for the same quarter one year earlier. As well as AWE and AEI the graph shows AWEWage. This is the wage component of AWE, calculated as the first term of the linear decomposition of equation (2.2). The decomposition is applied to monthly growth rates and not directly to the annual rate of change. The employment component and residual are discussed subsequently in section 4.4.

The employment term normally moves by under 0.2 percentage points. There is a peak of 0.6 percentage points in January 2007- again probably arising from a financial sector outlier. These changes persist in the rolling quarterly growth rate and thus headline growth of AWEWage is substantially below that of AWE for the whole of 2007. In this period, although not earlier, they account for most of the deviation between AWE and AEI.

We now proceed step by step to show how various changes bring AWE closer to the AEI. Figure 4-2 shows the growth of AWEWage with and without the adjustment for small firms, with the second series denoted AWES. This change does not have a great deal of impact.

Figure 4-2 AWE Wage with and without the Adjustment for Small Firms



In Figure 4-3 we show the impact of moving from the sample used in the AWE to the matched pairs used in the AEI with the second series denoted AWESM. The impact of this is substantial.

Figure 4-3 The AWE excluding small firms (AWES) and also using the AEI Matched Pair Sample (AWESM). Growth Rates of Rolling Quarterly Averages.

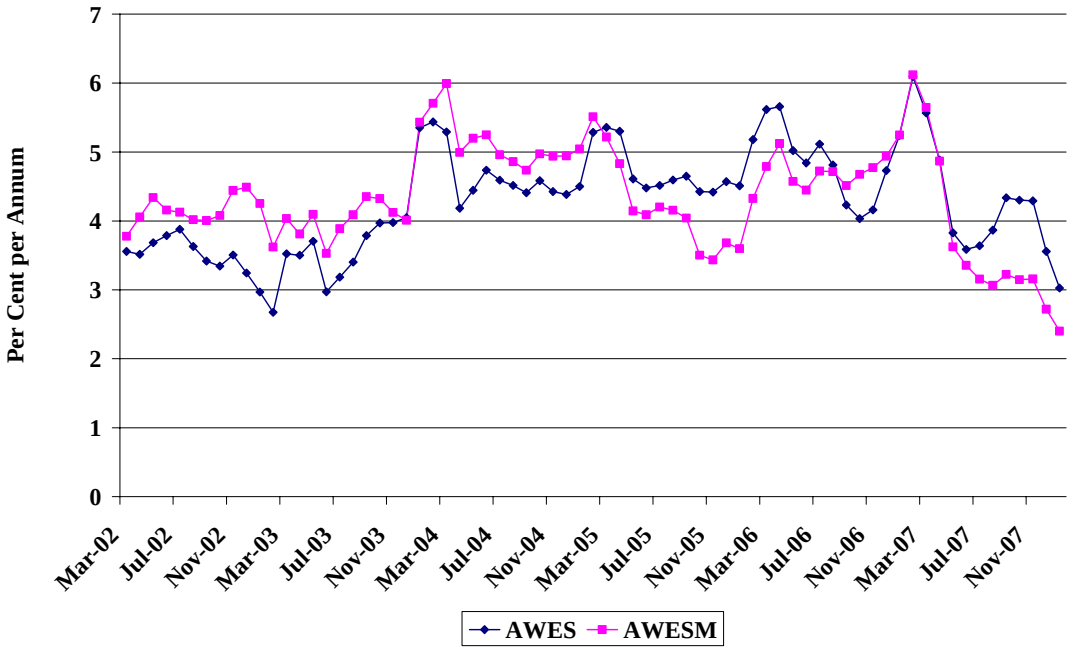


Figure 4-4 shows the impact of the classification differences which arise between the AWE and the AEI. In general the importance of this is not large but at the end of the period the effect is quite marked.

Figure 4-4 The Impact of Reclassifying AWE Firms on an AEI Basis

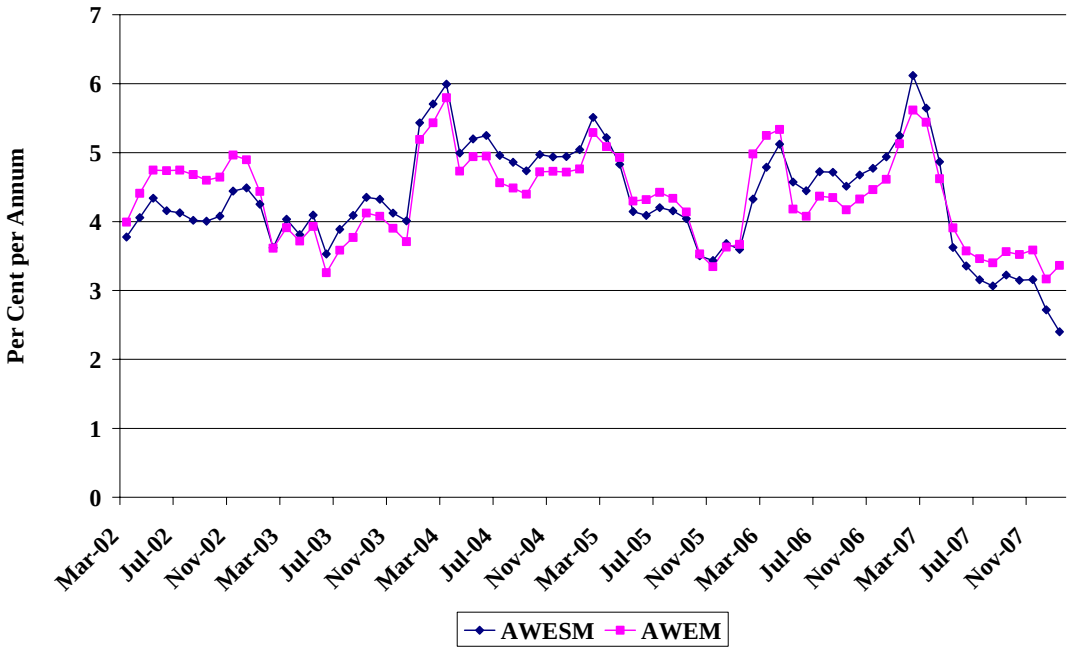


Figure 4-5 shows the impact of also introducing the AEI treatment of outliers. It should be noted that the AWE outlier approach works in terms of levels rather than growth rates, which makes it less likely to bias the growth rate substantially. The substantial impact in Figure 4-5 therefore reflects the AEI treatment which risks inducing a bias in the growth rate so as to achieve a substantial reduction in its variance. This effect is also visible in the differences between the means of AWEM and AWEIM in Table 4-1 although conventional statistical tests suggest that it is not statistically significant. The truncation process there is based on growth rates and it may be that more very large increases than very large decreases are being truncated. It should be noted that AWE distinguishes bonuses from other wages when identifying outliers. In effect it handles them separately.

Figure 4-5 The Growth Rates using Matched Pairs with the AWE (AWEM) and AEI (AWEIM) Treatments of Outliers. Growth Rates of Rolling Quarterly Averages

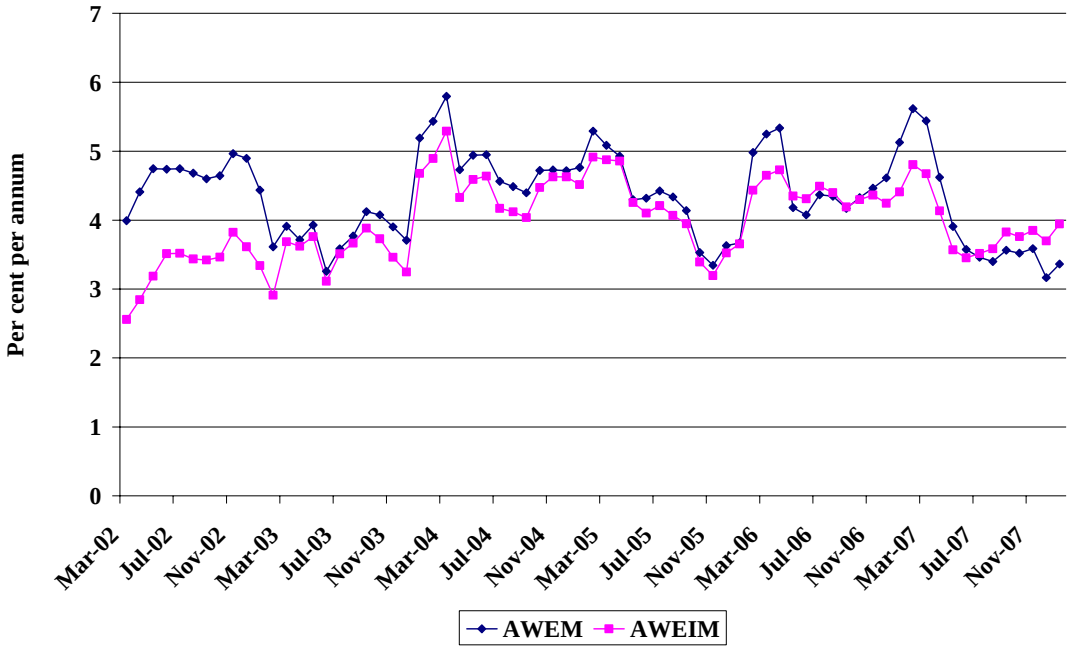
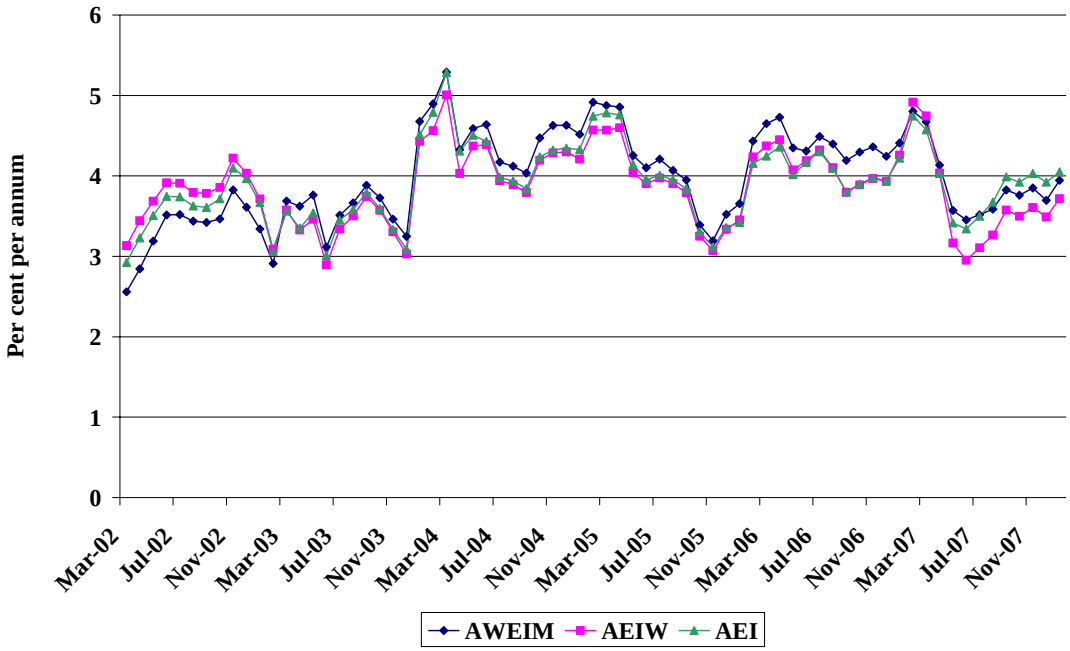


Figure 4-6 shows three series. The AWEIM uses employment weights which, since they are calculated from the MWSS, are as up to date as possible. The AEI uses fixed weights. However, the two different sets of weights are applied at a 116 industry level to estimates of the average wage in each industry. These average wage estimates differ because the estimates in each size-band are, in the AWE, combined using the latest information on the overall size of each stratum provided by the current IDBR, while in the AEI, the size of each stratum is estimated from the MWSS data and base-period IDBR data. Between these is AEIW. This is calculated using the fixed AEI weights to aggregate from the 116 industry to the 20 industry level and weights calculated from MWSS to combine the averages for these 20 industries. This suggests that the bulk of the effect of the variation in weights is observed in changing from MWSS weights and stratum size estimates to fixed weights at the detailed rather than at the broad level.

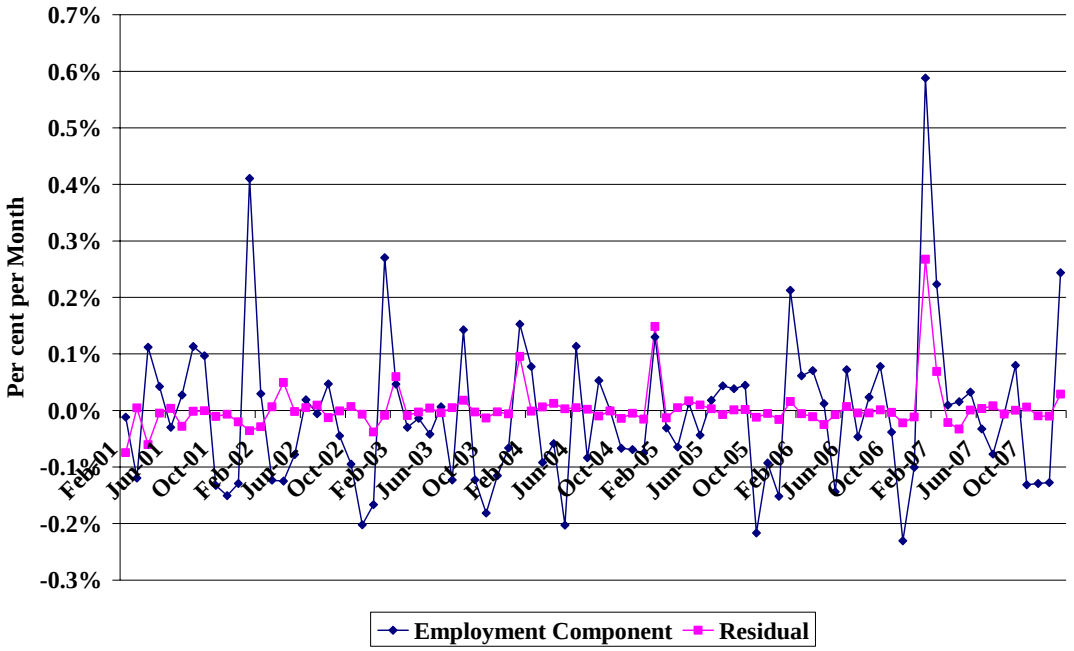
Figure 4-6 The AWE with Matched Pairs and AEI Outlier Treatment (AWEIM) compared with the AEI computed using AWE Weights (AEIW) and the AEI itself.



4.4 *The Employment Term and the Residual*

The move from AWE to AWEWage is accomplished using the linear decomposition of equation (2.2). Here we focus on the two other terms which emerge. The employment term shows the effect of the change to the employment structure on the aggregate measure, taking the wage rates as constant at the level in the first of the pair of months for which the growth rate is being worked out. We show first the monthly changes graphically in Figure 4-7.

Figure 4-7 Monthly Employment Components and Residuals in the AWE



While both the employment component and the residual are generally low and make no contribution of any importance to the overall growth of AWE, both plainly take outlying values in January 2007. In Table 4-3 we show the January means, both including and excluding the extreme 2007 figure. Even before this is taken into account it is clear that there is a strong January effect in the employment component; the 2007 case is simply an extreme example of this. On the other hand residual terms higher than 0.1 percentage points in absolute size are found only on three occasions.

Table 4-3 Key Statistics for the Employment Component and Residual

	Employment Component	Residual
Mean	-0.009	0.003
Mean for January 2001-2007	0.219	0.012
Mean for January 2001-2006	0.158	0.049

But at the same time we note that there was a substantial employment contribution in November 2006. There was a sharp fall in the share of Wholesale Trade which may have induced the large employment effect. This resulted from one firm being reclassified to Retail Trade. There was also a large contribution in May 2006. It should be remembered that, unless changes like this are offset by subsequent movements in the opposite direction, their impact will persist for fourteen months in the rolling quarterly comparison.

In other periods there have been substantial changes in employment shares which have had smaller impacts on the aggregate but are nevertheless of concern. For example in October 2005 the identified employment share of Public Administration fell to 3.99% of total employment from its value of 4.57% in September 2005 again as a result of reclassification.

If the primary focus of AWE is to provide an indicator of short-term growth then it is not desirable that substantial movements in the employment component and thus in the overall index should arise from reclassification. Suggestions for ways of dealing with the effects of reclassification follow in section 5.4.

4.5 Finding 3: The Differences between AWE and the AEI

The order in which the changes described above are made is to some extent arbitrary and influences the magnitudes of the effects attributed to them. Nevertheless, a number of inferences can be made independently of that.

First of all, comparison of AWE and AWEWage suggests that, except at the end of the period, the distinction between growth in the average wage and growth in the average before taking account of changes to employment mix, as shown by AWEWage, is not important. The difference in 2007 probably arises from labour market reclassification, and would be likely to be removed if recommendation 3 of Chapter 5 were adopted.

Similarly, the impact of the difference between the employment weights used in AWE and those used in AEI is explored in the comparison between AWEIM, AEIW and AEI. It is not particularly great most of the time although on some occasions it accounts for close to one percentage point of the difference between AWE and AEI. The major differences arise because of differences in treatments of outliers, missing observations and entry and exit into the sample, with matched pairs being suggested to reduce variability from the second and third of these points. For this reason the next chapter discusses the work which might be done so that ONS can be confident it is using the best practical approach.

These observations lead to the following finding:

ONS's decomposition of the differences between AWE and the AEI is coherent and provides an extensive assessment of the impacts of different features of the construction of the series on the differences between them. Nevertheless a slightly different decomposition has been presented in this chapter. A particular feature of this is that it separates the impact of changes in employment structure on the average wage into two components- those which arise from the effect of changes and those which arise because those changes have a subsequent impact on the way in which wages in individual industries are aggregated to produce an average. This decomposition makes it possible to identify the changes to the average wage which would occur in the absence of employment composition effects. This measure rather than a fixed-weight index should be used in order to understand the effects of wage changes within industries leaving aside the impact of changes in the mix of employment across industries (Recommendation 4).

Chapter 5. Recommendations on the Methodology of AWE

5.1 Introduction

Chapter 4 has suggested that the dominant causes of the discrepancy are the use of matched pairs, different treatment of outliers and the impacts of the different weighting schemes. At present no one is in a good position to say whether the AEI treatment of matched pairs and/or outliers is better or worse than that in AWE, given that the main aim of the index is to provide an indicator of wage growth. Thus the practical solution is to recommend the work which might be carried out to answer this question.

5.2 Research on Matched Pairs and Imputation

Background

The AEI is calculated using matched pairs to address the facts that either some firms might not reply in particular months or that disturbances may arise as some firms drop out of the sample and others replace them. In broad terms the matched pairs approach deals with missing observations by assuming that the rate of change of the average wage in the missing firms is the same as that in the observed firms. This approach was recommended in the Turnbull-King report.

Nevertheless, the recommendations in the Turnbull-King report were based on intuition rather than solid statistical evidence. Kokic and Jones (1998) undertook a study based on simulations of the ONS's Retail Sales Inquiry (RSI), to investigate the difference between matched pairs and ratio estimation. This work demonstrated that matched pairs estimates were more accurate when estimating changes over short periods than taking the difference between cross-sectional ratio estimators, but that this difference gradually eroded until around 15 months there was no difference, and for changes over longer periods the matched pairs estimators gradually drifted further away from the truth and hence were less efficient. It is not clear how much this 15-month result is related to the rotation period of the RSI sample, which was 15 months when this study was undertaken.

Much of this drift could be counteracted by benchmarking using a more accurate, annual estimate of the level, but there were two disadvantages. First, the act of benchmarking induces regular revisions in the series, and second, it is often necessary to wait 12-24 months (depending on the period considered) for the benchmark information to become available. One option to have some benchmark data available more quickly is to construct an annual estimate from the pooled monthly responses rather than waiting for an estimate from an annual survey.

The results were based on a simulated population which did not include any births or deaths - in a situation where births and deaths of businesses decrease the proportion of matched units, the performance of matched pairs might be expected to be worse. Smith, Pont & Jones (2003) used this work to argue that estimating changes from the difference in two cross-sectional estimates was a better strategy, although acknowledging that there were challenges with both this and the matched pairs

approach in practice. Cross-sectional estimation is certainly more robust to changing populations.

Issues Specific to Average Weekly Earnings

In AWE the matched pairs estimator which was used in the AEI has been replaced with separate cross-sectional estimates of earnings and employment; the ratio of these estimates forms AWE. This is largely as a reaction to the work summarised above. The headline figure for change in the AEI is the last 3 months compared with the same period a year ago, and this gives an average 12-month difference, which is only slightly shorter than the 15-month period over which the RSI simulation showed little difference between the two estimation approaches. On the other hand in the RSI simulation the matched pair estimator delivered better estimates of the growth rate over intervals shorter than 15 months. This might in itself create a presumption in favour of the use of the matched pairs estimator at least if the focus is to produce a good estimate of headline growth

It is not however clear what properties of the simulation give rise to this result, and it has been queried whether there is a strong foundation for moving away from the matched pairs used with the AEI. One relevant observation, however is that, while there is strong serial correlation between consecutive months' wages in the MWSS, that is also true of the RSI dataset.

There are several stages to the work which is needed to explore the relative merits of matched pairs relative to cross-sectional estimates of growth rates. These are:

1. Simulate values for the whole population of businesses for many consecutive periods.

This requires first that a model is built for the response values in the MWSS based on information which is available on the IDBR as predictors. Since variables covering both wages and employment must be modelled, there may be a need for a joint model in order to get a suitable simulation of the relationship between them. Then simulate values from this model, with appropriate random errors (probably by bootstrapping from the model residuals), for each business on the IDBR and for multiple time points based on the register information. (Although historical register information is stored it is difficult to extract so using actual changes from approximately eight years may not be practical. A first step might be to work with a stable population, only progressing to a varying population if there seems to be a need for further detail in the results). Kokic and Jones undertook this step only once, with simulation results conditional on the realised population; it would be possible to realise separate population values for each simulation and therefore to have more confidence that the solution is general, but may require more simulations.

2. Sample from the population according to MWSS design and rotation pattern.

This involves allocating permanent random numbers (PRNs) to each unit in the population, and then replicating the rotational sampling used in the MWSS sample design. Kokic and Jones varied the simulations from this stage, with new PRNs being allocated for each iteration.

3. Simulate non-response pattern

It is then necessary to construct a model of non-response based on the responses to the MWSS and register characteristics, and to simulate from this model in order to determine which of the businesses in the simulated sample are respondents.

4. Calculate estimators

Then calculate the matched pairs and cross-sectional estimators based on the responses, and use them to calculate the AWE.

5. Benchmark

Finally, benchmark the matched pairs series using the appropriate external information. There are three obvious alternatives - (a) the known population total from the simulated population, which will result in more accurate benchmarked estimates than can actually be obtained in practice; (b) an external benchmark from a further survey, calculated by appropriate sampling from the annual union of the simulated monthly data - this would require appropriate sampling and response mechanisms for the additional survey; (c) from the union of the monthly responses to MWSS, treated as a cross-section for estimation of the annual total - which has the advantage of being available around 12 months earlier than (b) if used in practice.

6. Summarise

Repeat steps 2 to 5 multiple times, then calculate statistics averaging over the simulations to assess the relative performance of each approach.

More general research on conditions for matched pairs estimation to work well

There might be a case for a more general analysis of matched pairs estimators. Under general conditions, the efficacy of matched pairs depends on the proportion of the sample which is matched between periods, which is affected by sample rotation and response patterns. The efficiency also depends on the increased correlation between periods of the matched pairs estimator relative to the correlation between consecutive cross-sectional estimators. These characteristics form a series of parameters which, taken with the overall variability of the survey data, may form a system which is either amenable to analytical solution or, more likely, can be studied by simulation from smaller and more standardised populations where these parameters can be manipulated directly. This may allow a more general solution to the conditions under which matched pairs estimators might work well, and then the characteristics of the AWE could be compared with this. This is likely to be a longer project, and one that will less certainly produce an answer to the specific problem of AWE.

Resources

Both of the approaches outlined above form substantial research projects. It is not trivial even to obtain the appropriate data and to produce models that adequately explain the patterns of response and non-response. .

Implications

The evidence such as it is at present points to a marginal preference for the use of matched pairs if the predominant function of the variable is to provide an indicator of

movements in average wages over a period of twelve months or fewer. The change from matched pairs to cross-sections is an important source of the difference between AEI and AWE. The terms of reference of this report draw attention to the need for a measure of short-term movements and, in view of this it does not seem that there is a good case for a change to a measure which, on the evidence that we have, performs better over an interval of more than fifteen months and worse over a shorter interval. However the use of the matched pairs estimator does make it difficult also to produce a good estimate of the level of average earnings. This issue is discussed further in section 5.6

This report therefore recommends that **AWE should not become a National Statistic until further work has been carried out on the possible use of matched pairs. This work needs to compare the use of matched pairs or a combination of imputation and matched pairs with the existing AWE methodology to see which produces more reliable estimates of annual growth rates. (Recommendation 1).**

5.3 Research to Establish Appropriate Treatment of Outliers

Background

The aim of outlier detection and treatment is to reduce variability of estimates even at the risk of introducing a small amount of bias. There is obviously a trade off between these two factors. The ONS standard approach is to use Winsorisation whereby parameters are set to minimise the mean squared error of estimates. The theory is well established and understood for the case of linear estimators

Lewis (2006) carried out an evaluation of outlier detection and treatment techniques which included a review of the literature pertaining to the theory of Winsorisation and implementation within three different National Statistics Institutes. One of the key papers referenced in this context was that by Chambers *et al*. (2000). He went on to propose and evaluate some options for Winsorising an estimator of change (in section 4). In particular, details are given (in section 5) of how the proposed methods were evaluated using real survey data from the Monthly Production Inquiry to form a pseudo universe with some simplifying assumptions regarding births and deaths.

Average Weekly Earnings

In AWE for practical reasons a simplistic outlier detection and treatment is currently utilised to reduce the mean squared error (MSE) of the level estimates of average wages at each survey period. The method is automated and contrasts with the largely manual methods employed within the AEI. The difference in estimators does not allow for a straightforward comparison of the two methods of detection and treatment (AEI only uses the matched sample) but nevertheless there is scope to build on Lewis' work in proposing a better methodology and evaluating against the current methodology employed within the AEI and AWE. The added complication of AWE is that the estimator for average wages at each time period is not linear but rather a ratio of two separate linear estimators for total wages and total employment.

Proposed work

An outline of a strategy for testing the current outlier methodology is given below.

1. Propose a suitable extension to the AEI methodology to include the non-matched sample such that it could be applied within the AWE estimation framework.
2. Carry out some theoretical analysis to develop a suitable method for Winsorising the AWE estimator to minimise the MSE for the change in average wage.
3. Test the methods developed in 1 and 2 against the existing AWE methodology using real MWSS data (see section 5.2.1 of Lewis, 2006).
4. Further test the performance of the three methods using a pseudo population of MWSS data (see section 5.2.2 of Lewis, 2006).
5. In the light of results of 3 and 4 propose a suitable and practical methodology for outlier detection and treatment for AWE which is fit for purpose.

Resources

The approach outlined above forms the basis of a substantial research project. The referenced paper was the result of a six month project though significantly includes a literature search. It is possible that the required work could be completed by using appropriate resources in about three months once those resources were identified.

Implications

While there is modest, but not statistically significant evidence that the AEI treatment of outliers leads to a bias in the long-term growths relative to the AWE treatment (page 31), it is not yet known whether the AWE treatment of outliers is an improvement on that of AEI. Therefore, this report recommends that **AWE should not become a National Statistic with the treatment of outliers present in the experimental series until it is clear that this is superior to the alternative which has been demonstrated to be workable in the construction of the AEI (but which would need development to be applied to AWE) (Recommendation 2).**

5.4 The Employment Component

The discussion above has focused on factors which explain why the wage component of AWE differs from that of the AEI. The employment component of AWE has no counterpart in AEI. There are nevertheless concerns about its erratic nature and its potential sensitivity to reclassification. The disturbances arise for the most part, but not universally, from the updating of the IDBR as the sample frame in January of each year. In turn these, as well as the changes discussed in section 4.4, are largely consequences of reclassification of firms from one sector to another.

ONS reports that a similar issue arises with the Index of Production. There the approach adopted is to link the index where the changes are substantial according to a threshold, so that, in the month when linking takes place, the total volume of production is estimated on both classifications. The volume index for the month in question is calculated using the figure from the old classification while the figure using the new classification is used to work forwards. Thus the growth rate is assumed to be uninfluenced by the reclassification. Appendix 4 provides further discussion of this issue.

The AWE figure is published as an amount (£ per week). If it were desired to maintain the focus on the level of earnings, then it would probably be best to taper the impact of classification changes both forwards and backwards, over a period whose choice would be bound to be arbitrary. The case for this is that a classification change is a step response to what is probably a smooth underlying process.

However the terms of reference of this review refer to the need to produce a measure of the short-term change of earnings rather than a measure of the level of earnings. Given this the process adopted with the Index of Production seems the appropriate choice, despite the fact that some focus on levels is also needed (see section 5.6). It is recommended that **the AWE series should be linked across substantial employment reclassifications so that the effects of reclassifications themselves do not affect the estimated rate of growth of the average wage (Recommendation 3).** Reclassification nevertheless has an impact because it leads to changes in the underlying weighting scheme of the type which create the difference between AEIW and AEI in Figure 4-6

5.5 A Fixed-Weight AWE

As the report has made clear, the AEI is calculated using fixed weights based on the composition of employment in the base year of the index. This review was asked "which AWE measure (fixed weights or current weights) should be used in which circumstance".

The motive behind the introduction of AWE is that, unlike the AEI, it measures the average wage in the economy. The AEI does not do this because it is not sensitive to changes in the structure of employment between industries.

Nevertheless, users may well be interested in knowing how much of the growth in wages is taking place because of changes to average wages within industries and how much is due to changes in the employment structure of the economy. The decomposition in equation (2.2) on page 15 addresses this. The ONS has produced a similar decomposition published as part of the AWE release. That decomposition differs from equation (2.2) because the residual is in effect allocated between what in equation (2.2) are the wage and the employment terms.

The ONS' decomposition does not answer the question "how much would the average wage have risen if the mix of employment had not changed?" while the wage term in equation (2) answers that question exactly. Neither of these, are however, fixed-weight indicators. The difference between them and fixed-weight measures is essentially that between a fixed weight Laspeyers index and a chain-linked index. It is fair to say that the merits of chain-linked indices over fixed-weight indices are generally recognized. Indeed the national accounts moved from a fixed-weight basis to a chain-linked basis in 1998.

In order to enable users to identify the effects of growth in wages within industries on the average wage it is recommended that **the ONS should publish, along with the main AWE results, figures showing the change to average earnings arising solely from changes to average wages by industry, and excluding any employment mix effect (Recommendation 4).**

In the experimental AWE the disaggregation above is presented for the whole economy and for the public and private sectors separately. It is expected that ONS would continue to provide the data on at least this level of detail. If there is user interest in a similar decomposition being carried out for broad industrial categories, it should also be possible to provide that. It is assumed that, as at present, ONS will continue to provide wage data including and excluding bonuses.

Identification of the wage effect raises the question whether ONS should also publish the employment effect as identified in equation (2). To do so would require it also to provide the residual or to indicate that the employment and wage effects do not add exactly to the overall change. Alternatively it could publish simply wage effects and other effects, where the latter include both the employment term and the residual. It should be noted that, following the changes recommended in section 5.4 the residual is unlikely to be more than 0.1 percentage points in any month but ONS may wish to wait until these changes have been implemented before coming to a final decision on the matter.

5.6 Levels or Growth Rates: the Weekly Wage or an Index

It should be noted that, if the main aim of AWE were to provide an indicator of the level of wages it seems unlikely that, given the current state of statistical knowledge, the approach adopted by ONS could be substantially improved on. A reasonable conclusion would be that the AWE methodology is already fit for purpose

However, as noted earlier, the terms of reference mention very clearly the need to produce a good measure of short-term changes in earnings and it is that need which results in the suggestions for further work made here. It also gives rise to a problem.

The treatment of employment reclassification proposed implies that it is not possible to produce an estimate of the level of wages; instead only an index can be produced. A similar issue arises if matched pairs are found to be a useful means of enhancing the quality of the indicator of wage growth. The employment classification issue could be addressed by tapering as suggested in Appendix 4, but it does not seem that the matched pairs issue could be resolved so easily.

As noted earlier, users also have an interest in estimates of the level of wages. Since there can be, and probably is in the context of measures of average earnings, a conflict between the production of good measures of the level of earnings and good measures of the growth of earnings this creates an inherent tension. One aspect of this issue is discussed formally in Appendix 3.1.

Over and above the points raised there, treatments of missing values and outliers which are desirable from the point of estimating growth well may induce biases. Even if biases are absent they may embed sampling errors relative to the underlying but unknown true value which are never corrected. In either case a measure of levels computed by cumulating a good measure of growth would tend to drift away from the true underlying level. This could be corrected by benchmarking the growth index to a level measure from time to time, but the benchmarking process would lead to revisions which users would be likely to find disconcerting, and ONS would have to explain to users that the revisions actually led to estimates of growth rates which were likely to be less reliable than its initial estimates.

The nature of the problem can be seen by considering the use of ASHE as a benchmark. ONS could, for example, calculate the AWE in growth terms and use it to roll forward benchmark wages as calculated from ASHE. Such an approach would be similar to that used elsewhere in the production of economic time series, with a benchmark rolled forward using an indicator variable. However, the difficulty arises when the figure for wages generated by rolling forward the wage calculated from ASHE for twelve months turns out to be rather different from that generated by the results for ASHE in the next year. As Table 3-1 shows, the growth rates derived from both AEI and AWE have differed from those derived from ASHE by up to 2 percentage points. It is difficult to see that the discrepancy would be very different if AWE were adapted to focus on growth rates rather than on levels. Thus users would have to accustom themselves to revisions which were large compared to annual growth rates if this approach were to be adopted.

One of the reasons why a good measure of growth might differ from a good measure of levels is that, as section 4.4 shows, the levels measure can be subject to abrupt changes as firms are reclassified. There may also be disturbances as some firms drop out of the sample and others enter. In both cases the impact on levels in any month would be smoothed out if the benchmarking were to take place not to any particular month but to average figures for say one year. This procedure has a clear parallel with many other ONS indicators which take as their base period the average for a year even though they are monthly data. However, there may be a case for rebenchmarking each month to the average of the last twelve months so as to avoid periodic sharp revisions. A scheme which may prove useful as a means of handling this is set out in Appendix 5. Alternatively it might be possible to benchmark AWE to national accounts figures for wages and salaries by industry. If this is done MWSS data will still be needed to handle the public/private split since this is not available in the national accounts.

Whatever approach is adopted it should reflect the following criteria:-

1. It should preserve long term growth in levels, to eliminate long term drift.
2. It should result in a series which is not sensitive in the short term to changes to the register and in particular to the effects of implementing the new register in January. Given also point 1 this means that register effects will be smoothed over a number of months.
3. The trade-off between revisions and the initial performance of the estimate of wage growth needs to be considered carefully in the light of user needs.
4. The method should be understandable to operational staff.
5. It should be easily applied at different levels of aggregation and ideally be additive.
6. It should be straightforward to programme.

It should be noted that if i) ONS decides to follow recommendation 3 or ii) it concludes that the use of matched pairs provides the best means of handling sample rotation, it will, nevertheless have to deploy methods very similar to those presently in use for calculating the level estimate of weekly earnings to be used to provide the benchmark.

It is therefore recommended that **ONS should research means of producing AWE so that it serves primarily as a short-term indicator of earnings growth but also delivers a coherent estimate of the level of wages (Recommendation 5).** The research methods are similar to those of the work on outliers and matched pairs, in that, by repeated sampling from MWSS data, ONS can explore the effects of different methods of benchmarking on both levels and growth rates of AWE. It will be able to identify the loss of performance of a growth indicator which results from benchmarking and conversely, the loss of performance of the levels indicator resulting from integrating it with the growth estimates.

5.7 Finding 4: ONS Methodology

As an indicator of the level of wages the AWE is close to fit for purpose. However, if the primary aim is to identify short-term changes to wages, further work is needed to establish whether the AWE methodology is better than the AEI methodology. (Recommendations 1 and 2). ONS has the clear capacity to carry out the required work effectively. There are also modifications which can be made without further research which should make the AWE more satisfactory (Recommendation 3). These have implications for the form in which AWE is presented and in turn imply that further work is needed to provide an indicator which meets the needs for both estimates of growth in earnings and of the level of earnings. (Recommendation 5).

Chapter 6. The Financial Services Industry

6.1 Introduction

The study was asked to examine the structure of the sample used for the MWSS and, in particular whether it gave enough attention to the financial sector. The figures for this sector have caused concern to some users and the AEI Statistician also reports that the data appear to have become more erratic.

6.2 Sample Allocation

The total budget for the MWSS determines the number of firms which can be approached. Since a census is taken of firms with 1000 or more employees, the size of that census in turn determines the number of firms with fewer than 1000 employees that can be sampled.

The theory of sample allocation is well understood. Cochran (1977, p.96) shows how an allocation can be chosen to minimise the variance of the key estimate when costs depend on the number of firms sampled and may vary from stratum to stratum. This approach forms the basis of past ONS work on the allocation of the sample, with the simplification that costs are not assumed to vary by stratum.

There are, nevertheless, important questions of judgement involved. Cochran's formula shows allocation depending on the variance of the population within each stratum; more firms are sampled from more variable strata in order to minimise the overall variance achieved with a given expenditure. But the variability of the strata is seasonal. The financial services sector is more variable during the bonus season than outside it. It is not normally thought practical to vary the sample size seasonally, because respondents need to know whether they are in the sample or not and are unlikely to understand seasonal participation. With the matched pair approach used in AEI seasonal variation of the sample is very difficult. Thus minimisation of the variance needs to be done taking a pragmatic view about the importance of data for different months.

A second problem arises from the tension between the reliability of the overall figure and the reliability of the figures for individual industries. An overall weighting scheme will give small sample allocations to strata which do not contribute very much to the overall figure. But if some users are interested in the data for these individual industries, then the sample allocation has to be higher. The ONS has to balance the competing interests of different users.

The last sample allocation exercise was carried out in 2003. A new sample allocation is needed in 2009 because the standard industrial classification is changing and the sample needs to adapt to the new classification. This obviously raises the question whether it is sensible to conduct a reallocation in 2008 as well. Given that the estimated cost of carrying out a full sample allocation is £15,000, ONS has proposed that as an interim measure the sample size for the financial sector be raised by 100 firms. This can be accommodated within existing budgets.

6.3 Conclusions and Recommendations

Since a full allocation exercise would inevitably divert resources from other areas of ONS work, this report recommends that, **as proposed by ONS, the sample size in the financial sector should be raised by 100 firms as soon as possible. The sample structure should be reassessed in 2009 in the light of the new Standard Industrial Classification (Recommendation 6).**

Chapter 7. AWE system requirements

7.1 Introduction

In the course of the review it became apparent that the software used in the production of AWE did not have the same capability as the AEI software. It does not appear that the AEI software is over-engineered; the statisticians compiling the AEI feel that it gives them a detailed understanding of the determinants of the eventual index values. There are a number of areas where the functionality of the AWE software falls short of that used to produce the AEI.

7.2 Revisions

ONS needs to have the ability to determine the main causes of the revisions to the previous months' data. Revisions can arise for a number of reasons – revised returned data, returned data replacing an imputation, returned data replacing data previously weighted for and changes to the estimation (the weighting or grossing) – and the total revision needs to be split into each of these categories and then within each category by contributor.

The analysis is needed by sector and 'industry'. In setting this up, ONS needs to think about whether this analysis should be done separately for earnings and employment. It would probably be helpful to distinguish the effects of outliers and outlier classification in this process. Some of the above analysis is currently available, provided via simple SAS programmes, but it is not complete and it is not integrated with the AWE system.

7.3 Changes

ONS needs to be able to determine the main causes of the differences between the current data and both the previous months' data and the data for the same month a year ago. The main changes will arise from growth or changes in contributor data, contributors previously weighted for or imputed but now returned (or vice versa) and genuine gains or losses arising from reclassifications or other movements out of the strata. A particular issue arises from the treatment of bonuses, if a firm is classed as an outlier in one year but not in the next year.

The analysis is also needed by sector and 'industry' and it is important to do it separately for earnings and employment. Again, some of this analysis is currently provided via simple SAS programmes but it is not complete or integrated.

One of the key features of the AEI system is the ability to 'drill down' and determine the main contributors driving the growth in the index. A similar system (a leverage system) is needed for AWE. There is currently a leverage system of sorts built in as part of the validation but each contributor is given equal weight (the system identifies contributors which are moving strongly month on month). The most suitable leverage measure is probably one which identifies the contribution of each individual firm to the change in the aggregate from one month to the next. This can be done by recalculating the aggregate excluding each observation, for each pair of months.

Where a firm appears in only one month, as is the case with firms joining or leaving the sample, then the firm is excluded only in the month in which it appears. ONS is confident that this calculation is practical.

7.4 Efficiency/Flexibility

The AWE system was designed to produce AWE on an experimental basis to the existing timetable, benefiting from much of the work done on the AEI. In comparison to the AEI round, there tends to be less intervention in the AWE round and there are far fewer runs during the results stages. In recent months the timescale for processing AWE has been moved forward, closer to the AEI and there have been more runs each month. It has become clear that the AWE system is less flexible than the AEI system.

This lower efficiency/flexibility may not be crucial but further testing is needed before AWE becomes a National Statistic.

7.5 Accuracy

The AWE was launched as an experimental series in August 2005. The programmes used to calculate it have not yet been independently verified. Given the importance of the series and the embarrassment that errors would cause, it is important that the quality of the computer programmes should be fully assured before AWE becomes a National Statistic.

7.6 Conclusions and Recommendations

Attention needs to be given to the software used in producing AWE. It is recommended that**, before AWE becomes a National Statistic, ONS should ensure that the software used to produce it has the same functionality as the AEI software with respect to revisions, changes and flexibility and, given the importance of wage statistics, that quality assurance of the programmes and procedures be carried out by a competent expert who was not involved with their development. The work should be done with an awareness of the needs of key users to understand what lies behind the movement in AWE (Recommendation 7).**

Chapter 8. Replacement of AEI by AWE

8.1 Introduction

Some of the differences between AWE and AEI exist for very clear reasons. Changes to the employment mix do not, by construction, change the AEI. But they do and should change a true measure of the average wage. Other differences exist because the ONS has, as a matter of choice, adopted an approach with AWE different from that used for AEI. Completion of the work recommended above will mean that, whatever choices the ONS finally makes about methodology, it will be able to say with confidence that, where this results in differences from AEI it does so for good reasons and that AWE is a superior measure.

8.2 Stakeholder Interest

In the course of this review ten probable users of the AEI were contacted in addition to the Bank of England and HM Treasury. A copy of the letter sent is attached in Appendix 6. No replies were received suggesting that users are not greatly interested in the issue. Nevertheless ONS cannot assume that this absence of interest will persist if AEI is discontinued.

8.3 Replacement Options

Logically there is no reason why, once the necessary work has been completed AWE should not replace AEI. ONS will certainly be able to calculate it back to 2001 and possibly back to 2000[5]. This will enable users to compare the movements in the two indicators, without needing further comparison in real time. The Bank of England has suggested that a longer back run of AWE data is desirable but conversely, recognises that , if this proves difficult or expensive to provide, it should not be an obstacle to the change.

However, it is likely that AEI is used for a wide range of purposes, and users will not start to think about the move from AWE to AEI until it is clear it is taking place, i.e. until after AWE has been approved as a National Statistic. This gives the ONS a number of choices:

1. Simply replace AEI by AWE and advise users to make any adjustments needed in the light of the overlapping historical series. This would have the advantage that there would be only one measure of wage growth so that, in this respect, confusion would be avoided. However the risk that it would cause more than mild inconvenience to users is high.

2. Retain AWE as an experimental statistic even though it has become a National Statistic, but announce that the change-over would be made in say twelve months time. This would have the advantage again that there is only one highlighted measure

[5] There are some differences between the AWE and AEI data sets in 2000 which are not fully understood and which may be an obstacle to taking the variant of AWE which is eventually adopted back to 2000.

of wage growth but the drawback that many users may still fail to appreciate that change is on its way.

3. Publish both AWE and AEI as National Statistics and announce that AEI will be discontinued in twelve months time. This would make it easier for users to plan for the changeover. However it would have the disadvantage that there would be two different National Statistics indicating movements in wages. There is a risk that the credibility of AWE would be damaged.

4. Publish the AWE as a National Statistic and make the AEI available only on the web as a down-graded statistic for a period of twelve months. This faces risks similar to 1, albeit in a milder form.

Of these options probably (3) is the most satisfactory despite the risks associated with it and also despite the fact that ONS will in effect be facilitating the use of an inferior statistic (AEI) when a superior statistic (AWE) is available.

8.4 Supporting Information

ONS will need to be in a good position to explain the change; it should find that this report provides helpful background material. At the same time the paper by Parkin (2008) will provide an additional and technical arithmetic explanation of the differences between the two series. However ONS will also need to provide clear public accounts of the results of the work proposed in sections 5.2, 5.3 and 5.6 so that users can be assured that the change has firm foundations in statistical research.

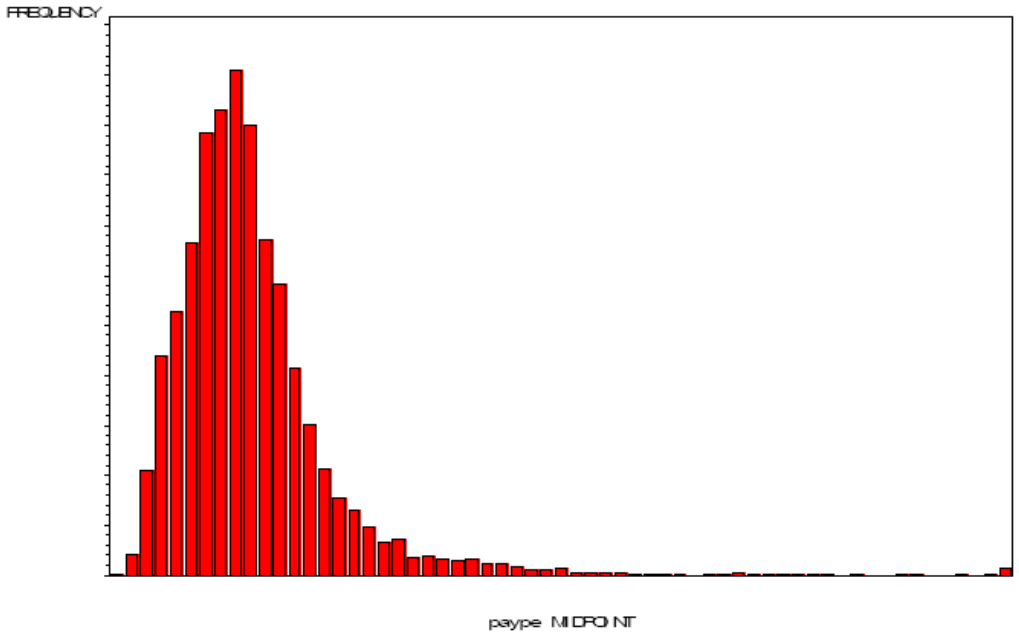
8.5 Conclusions and Recommendations

Given the logical superiority of AWE to AEI once the research questions of Chapter 5 have been resolved, **it is recommended that ONS replace the AEI with AWE once AWE has been given the status of a National Statistic but that both be published as National Statistics for a period of twelve months. Assuming that the parallel run does not produce any further issues, the AEI should then be discontinued. AWE data should also be made available back to at least 2001 and further back if possible.** (**Recommendation 8**).

Appendix 1: MWSS Data

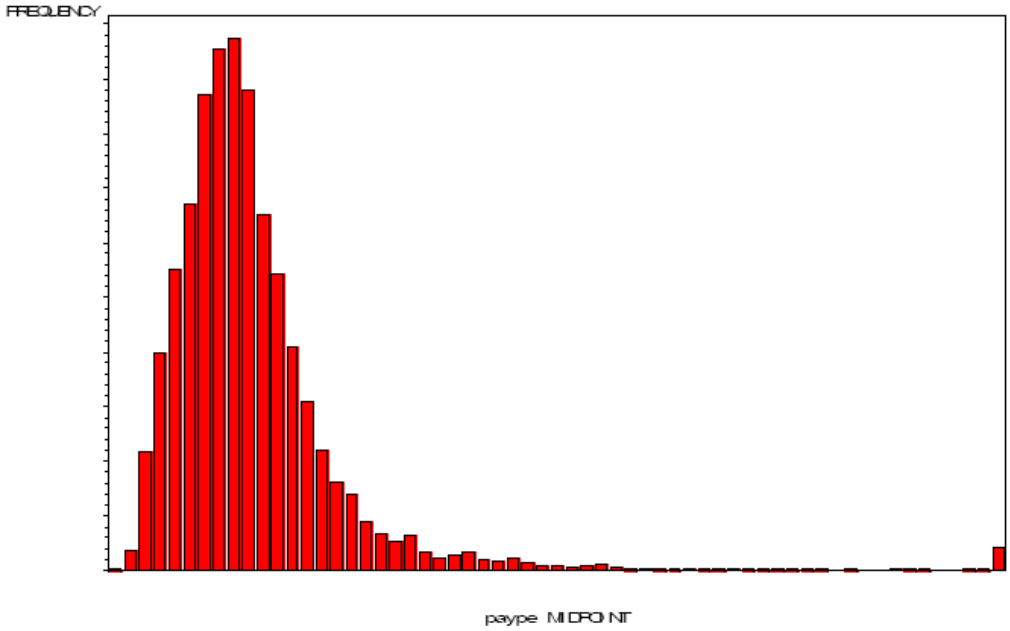
This appendix presents summaries of the MWSS data. The MWSS collects information on employment and pay. Bonuses and arrears are specifically identified. An indication of the nature of the responses is provided by figures A1-1 and A1-2. These show the distribution of average earnings per employee by employing unit. The first figure shows the data for November 2006, a month in which not many bonuses are paid. The second shows data for February 2007, a bonus month. Both figures show the skewness familiar from studies of individual earnings. In order to avoid the risk of disclosure of confidential information the graphs are drawn without scales.

Figure A1-0-1 The Distribution of Average Pay by Firm. November 2006



Average Weekly Earnings=£403

Figure A1-0-2 The Distribution of Average Pay by Firm. February 2007.

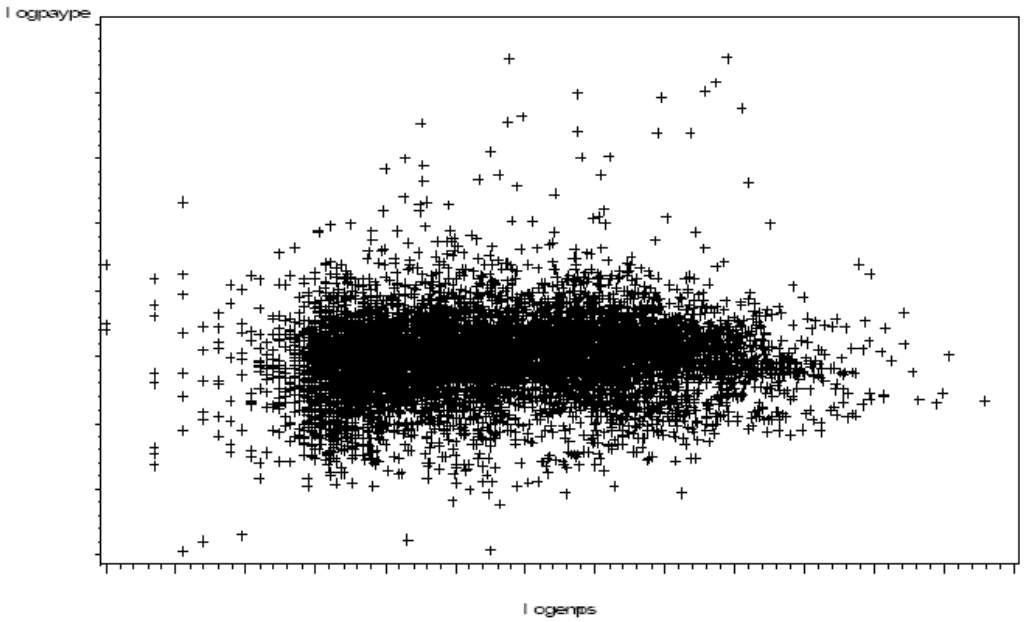


Average Weekly Earnings=£466.3

In the November figure the mean and the median are close. This is quite different from the experience of individual earnings and wage rates. By contrast, in February a relatively small number of firms pay large bonuses. This raises the mean substantially, but has little impact on the median.

While the survey does not allow us to say anything about wages in very small firms a comparison of the average wage in each firm with the number of employees does suggest that both very high average rates of pay and very low average rates of pay are more common in small firms than in large ones. This is illustrated in Figure A1-0-3.

Figure A1-0-3 Average Earnings and Firm Size (Log Scale). November 2006



Appendix 2: The Formulae for AEI and AWE

A2.1 Notation

It should be noted that there are four different relevant classifications used in the computation of variants of AWE and the AEI. These differences arise because:-
1. The industrial classification of AWE is revised to fall in line with the IDBR of the relevant January while the AEI classification is changed in August to reflect the IDBR of the previous January. This means that a firm can be in one stratum/industry in AWE and in another in AEI either because it is classified to a different industry or because its reported employment in the IDBR used for AWE is different from that used for AEI even though the industry is the same.
2. Use of different methods of identifying outliers in AWE and the AEI means that a firm can be classified to the main stratum in AEI and an outlier stratum in the AEI or vice versa. In AWE bonus outliers are distinguished from "core" outliers and this creates a further stratum.
3. Matched pairs are used in the AEI but not in AWE. This means that a firm which is present in AWE may simply be absent in the AEI with the implication that the samples for the two indicators are different even though the stratum/industry classification arising from IDBR may be the same in both cases.

We use the following indicators to denote the different classifications:

Table A2-1 Data Classifications

Variant	
E	AWE classification
ESM	Firms in AWE restricted to the matched sample used in the AEI
EM	ESM sample with reclassifications to the strata used in the AEI
EIM	EM sample with the AEI treatment of outliers
I	Sample used for AEI; this is the same as that used for EIM

In *E, ESM* and *EM* outliers are those identified using the six standard deviations definition adapted for AWE. Application of the six standard deviation rule means that as the classification changes some firms which might be outliers with one classification are no longer so with another classification or *vice versa*. However, we have used the outliers as identified in the *E* classification for *E, ESM* and *EM.*

h is the industrial classification for a stratum, defined by 2-digit standard industrial classification (SIC) combined with public/private classification. There are 116 public/private industries.

k represents the components of the 20-industry classification at which the AEI is published

g is the employment band/outlier stratum identifier. The identifier can take the following values: 0, 1, 2, 3, 4, 7, 8, 9. The meaning of these values is as follows:

0 for small business adjustment strata;

1 for 19 < employment < 100;

2 for 99 < employment <500;

3 for 499< employment <1000;

4 for 999< employment;

7 for AWE bonus outliers;

8 for type 5 AEI outliers; and

9 for type 6 AEI outliers together with pay and employment AWE type outliers.

i is the index of a reporting unit within a stratum.

t is the period to which the computations refer, for example January, 2006.

T is the fixed annual period to which certain quantities used in the AEI, such as the base period used in the calculation of employment weights, refer. Note that *T* is a function of *t*, that is *T=T(t)*, but this is always implied in the notation.

$U^{v,t}_{hkg}$ is the population of firms in stratum *g* of industry *h*as identified in the IDBR, with *v= E, ESM,EM, EIM* or *I*.

$S^{v,t}_{hkg}$ is the sample with *v= E, ESM,EM, EIM* or *I*.

With classifications *ESM, EM, EIM* and *I* which used matched pairs, the sample depends on whether one is looking at the first or second month of the comparison. S^{vt}_{hkg} refers to the sample used for the comparison of months *t-1* and *t.* $U^{v,t}_{hkg}= U^{v,t-1,t}_{hkg}$ and $S^{v,t}_{hkg}= S^{v,t-1,t}_{hkg}$ with *v=ESM, EM, EIM* and *I***.**

With classifications *EM, EIM* and *I* which used matched pairs, the sample depends on whether one is looking at the first or second month of the comparison. S^{vt}_{hkg} refers to the sample used for the comparison of months *t-1* and *t.* $U^{EIM,t}_{hkg}= U^{I,t}_{hkg}$, $S^{EIM,t}_{hkg}= S^{I,t}_{hkg}$, $U^{v,t}_{hkg}= U^{v,t-1,t}_{hkg}$ and $S^{v,t}_{hkg}= S^{v,t-1,t}_{hkg}$ with *v=EM,EIM* and *I***.**

N^{vt}_{hkg} For an ordinary stratum this is the universe stratum count in period *t*, minus the count of outliers in that stratum in period *t*. The outliers can be either AWE or AEI outliers. For an outlier stratum, it is the count of units in the outlier stratum. When matched pairs are used *t* relates to the second month of the pair for which the comparison is made.

N^{bvt}_{hkg} For an ordinary stratum this is the universe stratum count in period *t*, minus the count of bonus type AWE outliers in that stratum in period *t*, if AWE outliers are used. If the outliers used are AEI outliers then $N^{vt}_{hkg}= N^{bvt}_{hkg}$. For an outlier stratum, it is the count of units in the outlier stratum.

$n^{v,t}_{hkg}$ For an ordinary stratum this is the stratum sample size in period *t*, minus the count of outliers in the stratum in period *t*. The outliers can be either AWE or AEI outliers. For an outlier stratum, it is the count of units in the outlier stratum. The counts may be on the unmatched or matched sample, the context will make it clear which is implied. When matched pairs are used *t* relates to the second month of the pair for which the comparison is made.

$n^{b,v,t}{}_{hkg}$ For an ordinary stratum this is the stratum sample size in period t, minus the count of bonus type AWE outliers in the stratum in period t, if AWE outliers are used. If the outliers used are AEI outliers then $n^{t,v}{}_{hkg} = n^{b,t,v}{}_{hkg}$. For an outlier stratum, it is the count of units in the outlier stratum. The counts may be on the unmatched or matched sample, the context will make it clear which is implied.

$W^{v,t}{}_{hkgi}$ is the pay, including bonuses and arrears, recorded for unit *i* in stratum *hkg* in period *t* recorded on AWE dataset, when *v* equals *E*, or the AEI dataset, when *v=I*. Where firms report pay monthly, the monthly figures are multiplied by (7x12)/365.25. In cases *E* and *ESM* the figures for missing firms are imputed using the stratum growth rate from the last observation. There are slight differences between the AWE and the AEI files in some cases. The reasons for these are not clear but they are not of importance.

$B^{v,t}{}_{hkgi}$ is the bonus pay recorded for unit i in stratum *hkg* in period *t* recorded on AWE dataset, when *v* equals *E*, or the AEI dataset, when *v=I*.

$Z^{t}{}_{hkgi}$ is the employment for unit i in stratum *hkg* in period *t* recorded on the IDBR. In the AEI these are evaluated in period *T*, the base period of the index.

$E^{v,t}_{hkg} = \sum_{i \in S^{v,t}_{hkg}} E^{v,t}_{hkgi}$ is the stratum total for $S^{vt}{}_{hkg}$.

$\bar{E}^{v,t}_{hkg}$ is the average employment per firm in month t for sample $S^{vt}{}_{hkg}$.

$$\bar{E}^{v,t}_{hkg} = \frac{1}{n^{v,t}_{hkg}} \sum_{i \in S^{v,t}_{hkg}} E^{v,t}_{hkgi}$$

$\gamma^{vt}{}_{hkg}$ is the weight for ordinary pay and employment in AWE, for sample $S^{vt}{}_{hkg}$ in month *t*, accounting for all outliers

$$\gamma^{vt}_{hkg} = \frac{\sum_{i \in U^{v}_{hgi}} Z^{t}_{hkgi}}{\sum_{i \in S^{v}_{hgi}} Z^{t}_{hkgi}}$$

This adjusts for the possibility that the average size of the firms in the sample is different from that of the relevant stratum population.

$\gamma^{bEt}{}_{hkg}$ is the weight for bonus pay in AWE, for sample $S^{bEt}{}_{hkg}$ in month *t*, accounting for bonus type AWE outliers, when AWE outliers are used[6]

$$\gamma^{bEt}_{hkg} = \frac{\sum_{i \in U^{bEt}_{hg}} Z^{t}_{hkgi}}{\sum_{i \in S^{bEt}_{hg}} Z^{t}_{hkgi}};$$

when AEI outliers are used $\gamma^{bI,t}_{hkg} = \gamma^{I,t}_{hkg}$.

[6] For outliers the weights, $\gamma^{yt}{}_{hkg}$, are equal to one, and the stratum population size, $N^{vt}{}_{hkg}$, is equal to the stratum sample size, $n^{vt}{}_{hkg}$.

$$R_{hkg}^{v,t} = \frac{\sum_{i \in S_{hkg}^{vt}} \left(\left[W_{hkgi}^{v,t} - B_{hkgi}^{v,t} \right] + \frac{\gamma_{hkg}^{bvt}}{\gamma_{hkg}^{vt}} B_{hkg}^{v,t} \right)}{\sum_{i \in S_{hkg}^{vt}} E_{hkgi}^{v,t}} \quad \text{(A2.1)}$$

is the average pay per employee, including bonuses and arrears, for stratum *hkg* in period *t.* Note that if *v=EIM* or *I,* i.e. if the AEI treatment of outliers is used, bonuses are not distinguished from other pay so that

$$R_{hkg}^{v,t} = \frac{\sum_{i \in S_{hkg}^{vt}} W_{hkgi}^{v,t}}{\sum_{i \in S_{hkg}^{vt}} E_{hkgi}^{v,t}} \quad v = EIM \text{ or } I. \quad \text{(A2.2)}$$

$Z^{v,T}{}_{hk} = \sum_g Z_{hkg}^{v,T}$ is the total employment recorded on the IDBR for industry *h* in period *T* with definition *v*.

$Z^{v,T}{}_{hk0}$ is the total employment recorded on the IDBR for businesses employing fewer than 20 people, industry *hk* in period *t*.

$A^T{}_{hk0}$ is the factor for earnings for firms with fewer than 20 employees (the factors were computed from ASHE data, and are fixed).

Where matched pair estimation is used- in cases EM, EIM, IW and I- a single superscript *t* refers to the variable in time *t* when used in the second month of the comparision. *t-1,t* indicates the calculation for *t-1* when used in the comparison with month *t.* Thus $R^{I,t}$ is the average wage in month *t* when used in the comparison with moth *t-1* and $R^{I,t-1,t}$ is the figure for month *t-1* with which it is compared.

A2.2 The Average Earnings Index

$$R_{hk}^{I,t} = \frac{\sum_g N_{hkg}^{I,T} \bar{E}_{hkg}^{I,t} R_{hkg}^{I,t}}{\sum_g N_{hkg}^{I,T} \bar{E}_{hkg}^{I,t}} \quad \text{(A2.3)}$$

with the value for the previous month used in the comparison, denoted as $R^{I,t-1,t}$

$$R_{hk}^{I,t-1,t} = \frac{\sum_g N_{hkg}^{I,T} \bar{E}_{hkg}^{I,t} R_{hkg}^{I,t-1,t}}{\sum_g N_{hkg}^{I,T} \bar{E}_{hkg}^{I,t}} \quad \text{(A2.4)}$$

giving the analogous figure for the previous month computed for the same set of matched pairs.

Whole economy average pay is given by:

$$R^{I,t} = \frac{\sum_{h,k} R^{I,t}_{hk} Z^{I}_{hkT}}{\sum_{h,k} Z^{I}_{hkT}} \qquad \text{(A2.5)}$$

This value is based on the common sample between *t* and *t*-1. The equivalent value for period *t* -1 is calculated using the same common sample with month *t*. The growth in the index between *t*-1 and t is then given by the ratio of these two levels.

A2.3 Average Weekly Earnings

$R^{v,t}_{h}$ is an estimate of the average pay per employee in each industry/public private stratum.

$$R^{v,t}_{hk} = \frac{\sum_{g} \sum_{i \in S^{vt}_{hkg}} \left(\gamma^{v,t}_{hkg} \left[W^{v,t}_{hkgi} - B^{v,t}_{hkgi} \right] + \gamma^{v,b,t}_{hkg} B^{v,t}_{hkgi} \right)}{\sum_{g} \gamma^{v,t}_{hkg} E^{v,t}_{hkg}} \qquad \text{(A2.6)}$$

From these definitions we can now set out the calculations used to provide the various indicators shown in Chapter 4.

The AWE is computed according to the following formula,

$$R^{E,t} = \frac{\sum_{h,k} Z^{E,t}_{hk0} A^{T}_{hk0} R^{E,t}_{hk} + \sum_{h,k} \sum_{g} \gamma^{E,t}_{S_{hkg}} E^{E,t}_{hkg} R^{E,t}_{hkg}}{\sum_{h,k} Z^{E,t}_{hk0} + \sum_{h,k} \sum_{g} \gamma^{E,t}_{S_{hkg}} E^{E,t}_{hkg}} \qquad \text{(A2.7)}$$

It should be noted that the aggregate figure is calculated from the average wage in each industry. This is not given simply by aggregating the reported wages because the stratification of bonuses may be different from that of core wages. In particular bonus outliers are identified separately from core wage outliers.

A2.3 Other Variants of Indicators of Wage Growth

AWEWage

The subsequent measures look at the growth rate, and keep the employment composition unaltered. Thus

$$\begin{aligned} \frac{\Delta R^{EWage,t}}{R^{E,t-1}} &= \frac{\sum_{h,k} Z^{t-1}_{hk0} A^{T}_{hk0} \Delta R^{E,t}_{hk} + \sum_{h,k} \Delta R^{E,t}_{hk} \sum_{g} \gamma^{E,t-1}_{hkg} E^{E,t-1}_{hkg}}{\sum_{h,k} Z^{t-1}_{hk0} + \sum_{h,k} \sum_{g} \gamma^{E,t-1}_{hkg} E^{E,t-1}_{hkg}} / R^{E,t-1} \\ &= \frac{\sum_{h,k} Z^{t-1}_{hk0} A^{T}_{hk0} \Delta R^{E,t}_{hk} + \sum_{h,k} \Delta R^{E,t}_{hk} \sum_{g} \gamma^{E,t-1}_{hkg} E^{E,t-1}_{hkg}}{\sum_{h,k} Z^{E,t-1}_{hk0} A^{T}_{hk0} R^{E,t-1}_{hk} + \sum_{h,k} \sum_{g} \gamma^{E,t-1}_{S_{hkg}} E^{E,t-1}_{hkg} R^{E,t-1}_{hkg}} \end{aligned} \qquad \text{(A2.8)}$$

It should be noted that, taking the notation of equations (4.1) to (4.5),

$$\theta_{hk,t\text{-}1}^{AWE} = \frac{Z_{hk0}^{t-1} + \sum_g \gamma_{hkg}^{E,t-1} E_{hkg}^{E,t-1}}{\sum_{h,k}\left(Z_{hk0}^{t-1} + \sum_g \gamma_{hkg}^{E,t-1} E_{hkg}^{E,t-1} \right)}$$

AWES

The growth of AWES is calculated by leaving out the small firm correction from both the numerator and the denominator

$$\frac{\Delta R^{ES,t}}{R^{ES,t-1}} == \frac{\sum_{h,k} \Delta R_{hk}^{ES,t} \sum_g \gamma_{hkg}^{E,t-1} E_{hkg}^{E,t-1}}{\sum_{h,k}\sum_g \gamma_{hkg}^{E,t-1} E_{hkg}^{E,t-1} R_{hkg}^{E,t-1}} \text{ where} \qquad \text{(A2.9)}$$

$$R_{hk}^{ES,t} = \frac{\sum_g \gamma_{S_{hkg}}^{E,t} E_{hkg}^{E,t} R_{hkg}^{E,t}}{\sum_g \gamma_{S_{hkg}}^{E,t} E_{hkg}^{E,t}} \qquad \text{(A2.10)}$$

AWESM

For AWESM the firms in AWES are reclassified to the AEI classification but the calculations are the same as those for AWES

$$\frac{\Delta R^{ESM,t}}{R^{ESM,t-1,t}} = \frac{\sum_{h,k} \Delta R_{hk}^{ESM,t} \sum_g \gamma_{hkg}^{ESM,t-1,t} E_{hkg}^{ESM,t-1,t}}{\sum_{h,k}\sum_g \gamma_{hkg}^{ESM,t-1,t} E_{hkg}^{ESM,t-1,t} R_{hkg}^{ESM,t-1,t}} \text{ where} \qquad \text{(A2.11)}$$

$$R_{h,k}^{ESM,t} = \frac{\sum_g \gamma_{S_{hkg}}^{ESM,t} E_{hkg}^{ESM,t} R_{hkg}^{ESM,t}}{\sum_g \gamma_{S_{hkg}}^{ESM,t} E_{hkg}^{ESM,t}} \qquad \text{(A2.12)}$$

AWEM

AWEM is similar to AWESM but with firms reclassified to the strata used in the AEI

$$\frac{\Delta R^{EM,t}}{R^{EM,t-1,t}} = \frac{\sum_{h,k} \Delta R_{hk}^{EM,t} \sum_g \gamma_{hkg}^{EM,t-1,t} E_{hkg}^{EM,t-1,t}}{\sum_{h,k}\sum_g \gamma_{hkg}^{EM,t-1,t} E_{hkg}^{EM,t-1,t} R_{hkg}^{EM,t-1,t}} \text{ where} \qquad \text{(A2.13)}$$

$$R_{hk}^{EM,t} = \frac{\sum_g \gamma_{hkg}^{EM,t-1,t} E_{hkg}^{EM,t-1,t} R_{hkg}^{EM,t}}{\sum_g \gamma_{hkg}^{EM,t-1,t} E_{hkg}^{EM,t-1,t}} \qquad \text{(A2.14)}$$

AWEIM

AWEIM is structured in the same way as AWEM but with outliers identified as in the AEI rather than in AWE. This means that some firms which were place in outlier strata in AWE are not outliers in AEI and *vice versa*. The classification of firms in AWEIM is the same as that in AEI.

$$\frac{\Delta R^{EIM,t}}{R^{EIM,t-1,t}} = \frac{\sum_{h,k} \Delta R_{hk}^{EIM,t} \sum_{g} \gamma_{hkg}^{EIM,t-1,t} E_{hkg}^{EIM,t-1,t}}{\sum_{h,k} \sum_{g} \gamma_{hkg}^{EIM,t-1,t} E_{hkg}^{EIM,t-1,t} R_{hk}^{EIM,t-1,t}} \quad \text{where} \qquad \text{(A2.15)}$$

$$R_{h}^{EIM,t} = \frac{\sum_{g} \gamma_{hg}^{EIM,t-1,t} E_{hg}^{EIM,t-1,t} R_{hg}^{EIM,t}}{\sum_{g} \gamma_{hg}^{EIM,t-1,t} E_{hg}^{EIM,t-1,t}} \qquad \text{(A2.16)}$$

AEIW

Finally we come to AEIW, the AEI calculated with AWE employment rates

$$\frac{\Delta R^{IW,t}}{R^{IW,t,t-1}} = \frac{\sum_{k} \theta_{k}^{EIM,t-1,t} \sum_{h} \Delta R_{hk}^{I,t} \sum_{g} \left(N_{hkg}^{EIM,T} / n_{hkg}^{EIM,t-1,t} \right) E_{hkg}^{EIM,t-1,t}}{\sum_{k} \theta_{k}^{EIM,t-1,t} \sum_{h} R_{hk}^{I,t-1,t} \sum_{g} \left(N_{hkg}^{EIM,T} / n_{hkg}^{EIM,t-1,t} \right) E_{hkg}^{EIM,t-1,t}} \qquad \text{(A2.17)}$$

where $R_{hk}^{I,t}$ is given by equation (A2.3), and $\theta_{k}^{EIM,t-1,t} = \frac{\sum_{h} \sum_{g} \gamma_{hkg}^{EIM,t-1,t} E_{hkg}^{EIM,t-1,t}}{\sum_{k} \sum_{h} \sum_{g} \gamma_{hkg}^{EIM,t-1,t} E_{hkg}^{EIM,t-1,t}}$.

Notes

1. For outliers the weights, γ^{vt}_{hkg}, are equal to one, and the stratum population size, N^{vt}_{hkg}, is equal to the stratum sample size, n^{vt}_{hkg}.

2. In most months there are a small number of RUs used in the AEI that are not used in AWE.

3. In AWE the two extracts from the IDBR used to compute the stratum sample employment and stratum population employment do not match correctly, possibly because they are extracted at different times. The mismatch leads in turn the use of a makeshift adjustment to the g-factors. The author was surprised at the use of these adjustments when the preferred solution would be to extract the two sets of register counts in such a way that they do match. In the AEI a similar makeshift adjustment is made to the weights - suggesting that the stratum counts for that index also are not obtained at the same time.

Appendix 3: Levels and Growth Rates

A3.1 Basic Principles

There are important practical differences which can arise depending on whether a series is produced with the aim of indicating the level of a variable or its growth rate. The issue can be understood by looking at a variable *w* which is to be measured in period 1 and period 2. Does one want a good estimate of w_1 and w_2 or a good estimate of $(w_2 - w_1)/ w_1$. The statistician is often in a position to make choices about the way in which data are produced which should depend on whether users are mainly interested in the level or the growth rate.

Suppose $w_1 = w_1^* + \varepsilon_1$ and $w_2 = w_2^* + \varepsilon_2$ where the ε_i represent measurement errors with variance σ^2_i . Then the true estimates of the levels give the true estimate of the growth rate, *g**. But looking at what is actually constructed from the data

$$g = \frac{w_2^* - w_1^* + \varepsilon_2 - \varepsilon_1}{w_1^* + \varepsilon_1} = \frac{w_2^* - w_1^* + \varepsilon_2 - \varepsilon_1}{w_1^*(1 + \frac{\varepsilon_1}{w_1^*})} \approx \frac{w_2^* - w_1^* + \varepsilon_2 - \varepsilon_1}{w_1^*}(1 - \frac{\varepsilon_1}{w_1^*})$$

$$\approx g^*(1 - \frac{\varepsilon_1}{w_1^*}) + \frac{\varepsilon_2 - \varepsilon_1}{w_1^*}$$

To a first-order approximation, then *g* is an unbiased estimate of *g**. Its variance is given as

$$V(g) = \frac{1}{w_1^{*2}}\left(\sigma_1^2\left\{1 + g^*\right\}^2 + \sigma_2^2 - 2\left\{1 + g^*\right\}\sigma_{12}\right)$$

In many cases the variances and covariances are simply given. But where data collection is being designed the statistician is in a position to influence the relative magnitudes of the variances and co-variances. As a result choices have to be made about the dominant purpose of the data.

Suppose that, at each point in time *w* is evaluated afresh. Then, at least in a simple case there will be no covariance between the errors in the two estimates so that

$$V(g) = \frac{1}{w_1^{*2}}\left(\sigma_1^2\left\{1 + g^*\right\}^2 + \sigma_2^2\right)$$

For periods close together when the IDBR is not revised substantially, it is quite likely that the variances in the two periods are not very different from each other, so we can write

$$V(g_a) = \frac{\sigma_1^2}{w_1^{*2}}\left(\left\{1 + g^*\right\}^2 + 1\right)$$

But the statistician may be able to devise an alternative means of compiling the data where

$\varepsilon_2 = \varepsilon_1 + \eta_2$ with $\text{cov}(\varepsilon_1 \eta_2) = 0$

In this case $\sigma_2^2 = \sigma_1^2 + \sigma_\eta^2$ and $\sigma_{12} = \sigma_1^2$

$$V(g_b) = \frac{1}{w_1^{*2}}\left(\sigma_1^2\left\{1+g^*\right\}^2 + \sigma_1^2 + \sigma_\eta^2 - 2\left\{1+g^*\right\}\sigma_1^2\right)$$

$$= V(g_a) + \frac{1}{w_1^{*2}}\left(\sigma_\eta^2 - 2\left\{1+g^*\right\}\sigma_1^2\right)$$

so that

$V(g_b) < V(g_a)$ if $\sigma_\eta^2 < 2(1+g^*)\sigma_1^2$

Thus if the statistician can devise a method of computation with this property, the accuracy with which growth rates is measured can be improved on as compared to the growth rates which are estimated by a survey focused on levels.

It should be noted that whether this is the case or not is likely to depend on the interval over which the growth rate is calculated. Suppose that we are interested in measuring the growth rate, *g***, between period 1 and period *n* rather than simply from one period to the next. With the structure whereby measurement errors cumulate when method *b* is used, the condition becomes

$V(g_b) < V(g_a)$ if $(n-1)\sigma_\eta^2 < 2(1+g^{**})\sigma_1^2$

Whatever the value of the different variances, for some *n* large enough the growth rate will plainly be evaluated better using method *a* than method *b*. Thus, in choosing the method to be used in producing the data, the statistician needs to be aware not only of whether the users are concerned about levels or growth rates but also the interval over which the growth rate is to be calculated. Method *b* is subject to drift and eventually this drift offsets the benefits of its use for estimates of growth rates over short intervals.

In practice it may not be possible to identify an absolutely clear distinction between method *a* and method *b*. But the emphasis in AWE is to be as close to method *a* as possible. Subject to the sample in the MWSS rotating only slowly, the focus is entirely on calculating the level of the average wage. By contrast, the use of the data in the production of the AEI with its focus on matched samples and outliers defined with reference to growth rates is intended to resemble method *b.*

A3.2 A Composite Indicator

For a continuing survey with a rotating panel design, such as the Monthly Wages & Salaries Survey (MWSS), it is possible to estimate the monthly change directly from the difference between the two monthly estimates, made independently. However, a more accurate (smaller variance) estimate of the change can be made by taking advantage of the correlation between the responses for the same units (businesses in this case) in adjacent periods. Matched units are given greater weight because their change is more accurately estimated, and unmatched units are given a smaller weight because their change (which additionally includes a component of variance from the change of the units) is less accurately estimated. The weights for the optimum composite estimator depend on the relative size of the variance of the change in the matched and unmatched parts of the sample.

Practical issues

In principle it is possible to imagine joint composite estimation of (log) levels and growth rates in a fully consistent manner. However, this requires full knowledge of the covariance structures of the different estimates at all points in time and, so far as we are aware, there has not been any practical application of the method in this way.

The method has been applied focused on the question of estimating rates of change. Composite estimation used in this context involves adjusting the weights for estimating the *changes* between periods to make them more accurate. However, if the adjusted weights are used to estimate levels, they will be less accurate than if the unadjusted cross-sectional weights had been used.

Estimates of change formed using a composite estimator are not consistent with the difference in the estimates of the level. In fact the level at time t can be estimated in two different ways - one based on the weights from the t-1 to t change, and one based on the weights of the t to t+1 change, and these estimates will differ. Therefore in order to present the composite estimator, it will be necessary to publish a series of change estimates, without a corresponding series of level estimates (in order to avoid confusion from their inconsistency). This would be fine if interest focused on monthly rates of change but runs into problems when changes over longer periods are considered and particuarly when rolling quarterly figures are compared with those of a year earlier, as with the AEI. It is also a departure from standard ONS practice.

The inconsistency also appears between levels of aggregation, even within the changes. An optimum composite estimator for one sector will in general have different weights for the matched and unmatched parts (because of the different variances) from the optimum in a second sector. If we use these optima, then the weighted sum of the sectoral estimates will not equal the optimum overall estimate. If we use the weights for the overall estimate for all the sectors, then we do not have the best estimates of sectoral changes, and it is not clear whether these changes are better estimated than by taking the difference of two cross-sectional estimates. There is a possibility that the sectoral composite estimators could be calibrated to the overall composite estimate so that they are consistent, but this step would introduce additional variability, and it is less clear what concept is being measured by the calibrated estimates. (The same issues also arise for winsorisation.)

The use of a sequence of differences as the headline figure without the level would imply the need for levels to be given periodically. These would most naturally come from ASHE. There would then be a question of whether the changes, which act as a random walk about the true change and therefore potentially drift away from it, should be benchmarked to the annual change from ASHE. Such benchmarking would require annual revisions to the path. There are also some technical difficulties in that ASHE is a point-in-time estimate, and not an annual average, and so may be affected by changes in seasonality.

Conclusion

There is a range of issues which would need to be solved to implement publication of a wage series based on composite estimation in a satisfactory manner. To resolve these issues would require a substantial programme of research; this is unlikely to be a priority for ONS .

Appendix 4: Register updating for short-period surveys

ONS's business surveys use the Inter-departmental Business Register (IDBR) as their sampling frame. This is updated continuously with VAT information and quarterly with PAYE information, which between them update the classifications and sizes of some businesses (as well as providing information on new and defunct businesses). In a few cases these classification changes are real and reflect a change in a business's activity, but in many cases they are the result of correcting errors in coding. In order to (i) reduce the impact of these changes on the estimates of change between periods in the series and (ii) reduce the complexity of making adjustments for such changes, it was agreed in the mid-1990s that, for short-period (monthly and quarterly series) only, classification and size information would be fixed for a calendar year, and updated once at the end of the year (the "frozen register"). Although classification and size of existing businesses are fixed, note that there is no restriction on new businesses being added to the register, or on existing businesses being removed from the population when it is know that they have ceased to trade, so the population size remains responsive to changes in-year.

Exceptions are made for businesses with very large changes - either a change in 2-digit sector of the Standard Industrial Classification, or a change of size of more than 1000 employees - which are allowed to change within a year, and for which special adjustments may be made according to the circumstances of the change.

A4.1 Dealing with register changes at year-ends

Index numbers - linking

When the published statistic is an index, it is possible to make an adjustment so that the effect of changes which are artefacts is removed. The estimate for the three months October to December are calculated twice, once using the old, frozen register information from the year ending in December, and once using the newly frozen information for the year starting in January. If the ratio of these two three-month estimates is less than 0.95 or greater than 1.05 for a particular industry then the index for that industry is adjusted, with the ratio of the three-month estimates giving the adjustment ('link factor'). Conceptually this factor adjusts the series by multiplying each new period's estimate. Equivalently, however, it can be divided through the past data, and then the action of indexing to 100 reproduces the existing series until the present, and the adjusted series then joins on consistently without the need for further multiplication. Therefore the difference caused by register corrections is removed. In the most recent years no such adjustments have been needed.

Value series - wedging

Value series are generally dealt with differently. It would be perfectly possible to link them as described above, but this would require repeated revisions to the historic level of the series, in order for it to remain consistent with the most recent register information. This is generally not helpful, because the linking eventually removes most of the sample-based evidence for what the estimated total should have been. Instead in these cases differences can either be accepted as part of the noise of the series, or wedged in - that is, the difference is divided into *n* pieces, and a piece is

added to the series each period for *n* periods so that the new level is reached. Where values series are published from the same source as the Index of Production (that is, for engineering sales and orders), the current approach is to accept the differences.

Consistency

The difference in approaches for index numbers and for values means that these series are not consistent. The Retail Sales Index does have consistency between index and value because neither the index nor the value series is adjusted except occasionally in exceptional circumstances. But the Index of Production and Index of Services are potentially adjusted, which could lead to inconsistencies between the index and value series at year ends.

Appendix 5: A Benchmarked Estimator of Wage Levels

Suppose that in month *t* there is an estimate of level of wages computed in a manner similar to that currently used for AWE and denoted w_t. There is also an estimate of the growth rate which differs because i) linking is employed to smooth out the effects of employment reclassifications and or because ii) matched pairs are used to smooth out the effect of sample rotation. We denote this estimate of the growth rate g_t. If the growth rate were calculated from the level variable then we would have

$$g_t = \frac{\Delta w_t}{w_{t-1}}$$

and users would expect this to be a property of any published data series.

One possible approach to addressing the issue would be as follows:-

Construct a moving average estimate of the level over the last twelve months

$$\bar{w}_{t-5.5} = \frac{\sum_{i=0}^{11} w_{t-i}}{12}$$

Use the growth rates to give monthly figures, w^*_t taking account of the fact that the mean is centred between months *t-5* and *t-6*

$$w^*_{t-5,t} = \bar{w}_{t-5.5} \times (1+g_{t-5})^{1/2}; \quad w^*_{t-k,t} = \bar{w}_{t-5.5} \times (1+g_{t-5})^{1/2} \prod_{i=k}^{4} (1+g_{t-i}); \quad 0 \le k < 5$$

If the same method is applied, month by month to each published datum, then the growth rate between months *t-1* and *t* derived from the published series will be

$$g^*_t = \frac{w^*_{t,t}}{w^*_{t-1,t-1}} - 1 = \frac{\bar{w}_{t-5.5}}{\bar{w}_{t-6.5}} \frac{(1+g_t)}{(1+g_{t-6})^{1/2}(1+g_{t-5})^{1/2}} - 1$$

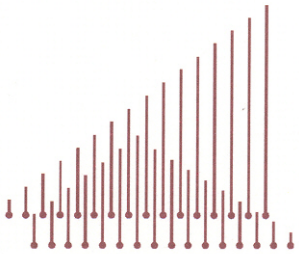
Any persistent disturbance to the level term will take twelve months to build up to its full impact.

If ONS is to meet the request from the Bank of England and HM Treasury for series which show the level of wages as well as their short-term growth, then it should assess the performance of estimators of this type. The key criteria it should examine are

i) the difference between g_t and g^*_t and
ii) the revision pattern of w^*.

It is not possible to give firm guidance here as to what balance between these two is either acceptable or most desirable. Having examined one or more benchmarking methods in this way ONS should come to a conclusion in the light of discussions with key users.

Appendix 6: Letter sent to Stakeholders



National Institute
of Economic and
Social Research

2 Dean Trench Street
Smith Square
London SW1P 3HE
Telephone +44 (0) 20 7222 7665
Facsimile +44 (0) 20 7654 1900
e-mail enquiries@niesr.ac.uk
http://www.niesr.ac.uk

14th February 2008

Dear

On behalf of the Office for National Statistics and with the support the Bank of England and HM Treasury, I am conducting a review of the measure of Average Weekly Earnings (AWE). The review was announced yesterday by the Office for National Statistics (http://www.statistics.gov.uk/pdfdir/awe0208.pdf).

As you may know, the AWE is intended to provide an up to date measure of average weekly earnings per person employed in the economy, while the Average Earnings Index (AEI) is closer to an index of wage rates. From time to time the rate of growth of AWE, which is still an experimental statistic, has departed substantially from that of the AEI.

The terms of reference of the review include

- whether the ONS's proposed methodology for AWE is fit for purpose for the principal measure of short-term changes in earnings;
- whether AWE is a better measure of short-term changes in earnings than the AEI;
- which AWE measure (fixed weights or current weights) should be used in which circumstance;
- the nature and extent of supporting information required to supplement AWE;

- whether, after a transitional period, the AEI can be withdrawn.

I would welcome your views, as a user, on any of these points. In particular I would like to know whether, for your purposes, you find a measure of average weekly earnings (assuming it represents this reasonably accurately) more or less useful than the current AEI.

Assuming that the AWE does its intended job satisfactorily and it becomes a national instead of an experimental statistic, it would be possible to publish both current weight measures (like the present AWE) and fixed-weight measures (closer to the AEI conceptually but still rather different sometimes in practice). Would you expect to find this helpful or a source of confusion?

In my report I intend to provide a list of people consulted; it may also be helpful to attribute views to individuals. Please could you therefore advise me whether i) you wish any comments you make not to be attributed to you and ii) whether you are happy for your name to be included in the list of people consulted.

In order to allow me to follow up points where that seems helpful, I would be grateful for a reply by 7th March 2008.

Yours sincerely,

Dr M.R.Weale.

References

Chambers, R.L., P.N. Kokic, P Smith and M. Cruddas. (2000) Winsorisation for identifying and treating outliers in business surveys. *Proceedings of the Second International Conference on Establishment Surveys (ICES II)*, pp 687-696.

Cochran, W.G. (1977). *Sampling Techniques*. Wiley New York

Kokic, P. and T. Jones. (1998) Comparing estimation methods for a monthly business inquiry. *Proceedings of Statistics Canada Symposium 97 - New Directions in Surveys and Censuses*, 269-272. Statistics Canada, Ottawa.

Lewis, D.J. (2006) An examination of the use of Winsorisation in sample surveys, Masters Thesis, Division of Social Statistics, University of Southampton.

Parkin, N. (2008). A Comparison of the Average Earnings Index and Average Weekly Earnings. Office for National Statistics.

Smith, P., M. Pont and T. Jones. (2003) Developments in business survey methodology in the Office for National Statistics, 1994-2000 (with discussion). *Journal of the Royal Statistical Society, Series D* 52 257-295

Turnbull, A. and M. King. (1999). *Review of the Revisions to the Average Earnings Index*. House of Commons. HC263.