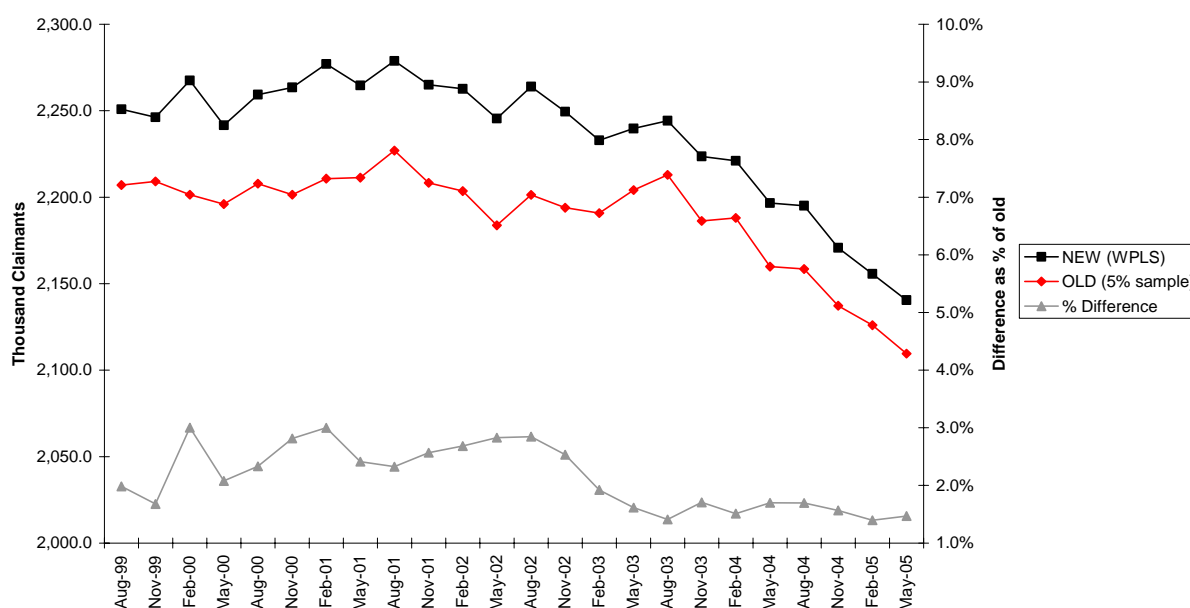


Income Support (IS)

Work and Pensions Longitudinal Study (WPLS) data compared to currently published 5% sample data.

IS comparison of caseloads (excluding MIG), 5% sample and WPLS data



The graph above shows that WPLS Income Support data (excluding MIG cases) is, on average, 47,000 or 2.1% higher than the 5% sample data.

The figures from the WPLS data appear better than the 5% sample data because they are more complete, capturing more backdated claims (retrospection) that were live at the enquiry date (quarter).

The difference between the WPLS data and 5% sample data is explained by:

1. **Retrospection.** Retrospection explains most of the difference, resulting in the WPLS data including on average an additional 40 thousand cases new claims that are missing from the 5% sample data. (Further information on retrospection can be found in the Glossary below)
2. **End dates.** On the WPLS data scans exact end dates are not received therefore they are estimated. This means, on average, around an additional 14,000 cases may be counted as live on the WPLS when they have very recently terminated.
3. **Cleaning.** The 5% sample data has some cases removed by use of automatic and manual (manual - prior to Feb05 only) cleaning rules. This removes some claims that are not live and in payment. The WPLS data is not able to be cleaned in the same way so includes, on average, approximately 4,000 additional cases.
4. **Clerical cases** – Clerical cases (paper records of claims not yet updated onto the computer system) are included on the 5% sample data (prior to Feb 05 only) but not on the WPLS data. Therefore, prior to Feb 2005, the 5% data will include an additional 4,000 cases.
5. **Sampling error** may explain a difference in totals between the WPLS and 5% of up to about +/- 13 thousand. This is the 95% Confidence Interval around the 5% sample figure.

Glossary

Retrospection arises from the delays which occur in practice between some new claims becoming eligible and their entry onto the computer system. If a case has not been added to the computer system by the time the 5% sample data is extracted (the reference date), it will not be counted (although benefit will be paid). The WPLS, however, incorporates information added to the computer system after the reference date.

The addition of such 'late' information is an important source of difference between the caseloads obtained from 5% sample and WPLS data, and generally results in the WPLS including more cases than the 5% sample data.

Late terminations

Delays can exist between a claim terminating and its removal from the computer system. Such cases will be included in the caseload obtained from the 5% sample data. When information is subsequently received that the claim has ended through claimant death, these cases are removed from the 100% data with respect to a point in time. However, any other form of late termination is not accounted for in the WPLS.

Sampling error

By definition, WPLS data yields the true population value. By contrast statistics published from 5% sample data (obtained by rating up by a factor of 20) are subject to sampling error and are actually estimates of the true population value. By chance, an estimate of the population value, obtained from the 5% sample data, may be slightly lower or slightly higher than the true population value.