

Wage and Employment Dynamics

Research Development Workshops

21 November 2022, Sheffield: University of Sheffield
2 December 2022, London: Bank of England

Wage and Employment Dynamics (WED)

WED datasets currently available

Enriched ASHE

The Annual Survey of Hours and Earnings (ASHE) is carried out in Spring each year by the UK Office for National Statistics (ONS). It is intended to be a 1% sample of employees, with data provided by employers. The same individuals are sampled in April each year that they hold an employee job, meaning their longitudinal employment path can be studied.

As data is supplied by employers, the information is limited to what can be supplied from payroll records: detailed information on wages and paid hours, employer pension contributions, occupation, industry and location, but only age and gender for personal characteristics. From the survey form and its own records, ONS generates several useful derived variables.

The data can be linked to other ONS business datasets that contain data on the employer (the enterprise) through the Inter-departmental Business Register (IDBR) Enterprise Reference numbers that are provided on the ASHE dataset. However, workplaces can only be indirectly linked to workplace data in other datasets as ASHE uses its own reference system for workplaces.

ASHE data is available from 2004 onwards. Its predecessor, the New Earnings Survey (NES) ran from 1975; it used the same sampling frame and collected much the same variables, so potentially information is available on some workers for over 40 years. There are some concerns about linking individuals over this period, but the 'ASHE' data files supplied to researchers by ONS include NES data from 1998-2003 which has been reliably linked.

ASHE is generally seen as high-quality data as it is provided by employers from payroll records rather than relying on employee recall. Large employers directly supply ONS with payroll records, rather than completing individual forms for members of staff. A major potential issue with ASHE is the sampling and weighting - see WED website from more information:

<http://www.wagedynamics.com/research-2/>

ASHE is point-in-time information. There is no information on what ASHE respondents do through the rest of the year, while employees in the ASHE population but not working during the April reference week are not included in the dataset.

The standard ASHE dataset has been available to researchers through the ONS's Secure Research Service (SRS) and the UK Data Archive Secure Data Service for many years, as well as being directly distributed to government departments. The WED team have generated an enriched dataset (Enriched ASHE), which includes additional variables derived by the WED team. The key difference

Wage and Employment Dynamics

Research Development Workshops

21 November 2022, Sheffield: University of Sheffield

2 December 2022, London: Bank of England

between the Enriched ASHE and standard ASHE datasets is the creation of additional variables by running Stata code available to all approved researchers¹.

ASHE does not cover NI; data is collected by NISRA on a consistent basis, but this is not generally available to GB researchers.

For more information, please refer to the Enriched ASHE Quick Start Guide which is available on the WED website:

<https://www.wagedynamics.com/wp-content/uploads/2022/09/ASHE-Quick-Start-Guide-v01k.pdf>

ASHE – 2011 Census

One drawback of the Annual Survey of Earnings and Hours (ASHE) dataset is that it contains relatively few personal characteristics for the employee (restricted to gender, age and residential location).

In order to expand the number of personal characteristics that are observed for employees in ASHE, ONS has sought to link the personal details of employees observed in the 2011 ASHE to those of individuals observed in the 2011 Censuses for England and Wales (CEW11). The linking process proceeded in two phases, described as 'Phase 1' and 'Phase 2'. Data from Phase 1 is currently available for researchers to access, while Phase 2 should be available for users to access by the end of 2022.

In Phase 1 of the linking process, robust links have been made for 62% of the employee job records in the 2011 ASHE belonging to employee's resident in England and Wales. In many cases where a robust link has not been made, this is because missing data on one or more of the core linkage variables (name, gender, age or residential postcode) has prohibited the reliable identification of the ASHE employee in the Census dataset. Phase 2 of the linking process brings in additional linkage variables in order to achieve robust matches for some of these cases. The linkage rate will increase by up to 10 percentage points after Phase 2.

In cases where a robust link has been made under Phase 1, it has been possible for ASHE (the "recipient dataset") to incorporate employee-level data from CEW11 (the "donor dataset") on a range of personal characteristics, including the employee's educational qualifications, country of birth, ethnicity, religion, and disability status. It has also been possible to incorporate some limited information about their household circumstances. These Census data items are available for the matched individual in any year that they appear in ASHE from 1999-2018.

¹ The code is available to researchers in the SRS, along with additional documentation. There is an additional data set generated by ONS with some extra variables (WED-ASHE). This dataset is not currently available for general use by researchers but is expected to be released in December 2022

Wage and Employment Dynamics

Research Development Workshops

21 November 2022, Sheffield: University of Sheffield
2 December 2022, London: Bank of England

A current version of the ASHE-CEW11 User Guide is available via the WED website and provides researchers with an overview of the linkage process and linkage outcomes from Phase 1. It also describes the weights that have been constructed to account for linkage biases, and it explains how potential users can gain access to Drop 1 of the ASHE-CEW11 dataset. Once Drop 2 is made available at the end of 2022, the user guide and supporting metadata listing the new variables will also be updated and uploaded to the WED website.

The linked ASHE-CEW11 dataset is being made available to researchers via the ONS Secure Research Service and provides opportunities for new research on issues such as the returns to education and wage inequality.

In a separate exercise, the Northern Ireland Statistics and Research Agency (NISRA) has linked the ASHE data for Northern Ireland to the 2011 Census for Northern Ireland. These data are made available by NISRA.

ASHE – 2011 Census quick Start Guide is available from:

<http://www.wagedynamics.com/wp-content/uploads/2022/09/ASHE-CEW11-User-Guide-Version-1.2-Drop-1-1.pdf>

WED datasets to be made available over the next 12 to 18 months

ASHE – PAYE linked dataset and ASHE – Self Assessment linked dataset

The WED team are working with the ONS to create a fully documented linked administrative dataset using ONS's ASHE dataset and linking it to HM Revenue & Customs' (HMRC) PAYE RTI and self-assessment data sets. The dataset should be available for accredited researchers to access via the ONS's Secure Research Service by the end of 2023.

As ASHE is an annual survey, the linked PAYE data should benefit researchers by filling in gaps in individual working profiles. For example, the inclusion of PAYE data will fill in gaps on labour market transitions that occur in the 12 months between ASHE survey intervals and enable the validation and calibration of earnings survey data. While HMRC Self-Assessment (SA) data will provide a broader picture of an employee's income beyond earnings, and the inclusion of SA data from the self-employed will expand the view beyond employee jobs.

ASHE _ Migrant Workers Scan

The WED team area also working with ONS to link HMRC's data on migrant workers to the ASHE data. The linked dataset will make it possible to better understand the working experiences of migrants compared with the rest of the population in Great Britain. The dataset should be available for accredited researchers to access via the ONS's Secure Research Service at the beginning of 2024

Wage and Employment Dynamics

Research Development Workshops

21 November 2022, Sheffield: University of Sheffield
2 December 2022, London: Bank of England

The HMRC Migrant Workers Scan contains information on all overseas nationals who have registered for and been allocated a National Insurance Number. As migrants can be tracked over time, it will support a dynamic analysis of migrants' progress through employment. Key questions the dataset could help to answer include:

- What role do migrants play in the labour market?
- What is the impact of migrants on the labour market?
- How do migrant workers earnings progress through their career?
- What role do migrants play in supporting particular sectors, such as public services?
- How do migrants' employment experiences based on gender, age and residence compare to non-migrants