

QA of sjd ('same job' marker) and sernol (serial number in previous year)

Document not yet cleared for intermediate/ publication release

One-page summary

Issue/area of concern

The variable 'sjd' reports whether a person has been in the same job (note, not 'same employer') for 12 months or more. This is a survey question, and valuable for research into within-work progression. If marked 'yes', the understanding (in the 2000s) was that information from the previous year was copied over, but this is not currently the ONS understanding. How sjd is used and how it relates to the independence of other variables therefore has the potential to affect analysis.

In the new WED drop, a variable 'serdol' was included, indicating the reference number of the same job in the previous year. This is highly valuable, as it allows reliable longitudinal jobs to be constructed. It is not known yet how this variable is constructed, or how well it represents job continuity.

Findings

Serdol was used to construct a persistent job identifier for an individual, 'job_number'. Job_number and sjd were analysed for paired years (2009/10, 2013/14, 2017/8) to see whether the longitudinal link seemed reliable. Sjd could only be analysed for an employee with a single job in both years. Job_number was analysed for all jobs held for two years in a row as identified by serdol.

Overall, the results are consistent with both variables correctly identifying the same job. However, the construction and use of these variables still needs to be clarified

Changes made to data

In the WED dataset, we have created a new variable job_number – a reference number for a continuing job, unique within the individual. 8-digit number.

Recommendations to researchers

At present both sjd and job_number seem reliable for constructing longitudinal job links. Note that sjd does not distinguish between jobs if an individual has multiple jobs in one year.

Details of QA

Analysts: Arusha McKenzie, Felix Ritchie

QA Code file: Main job analysis, P:\Working\code\sub_qa_main_job_analysis_v001.do

Variable creation code: can be found in the main ASHE creation code

This version: August 2021

Detailed analysis

Summary:

- The variable 'sjd' is supposed to reflect person is doing the same job for the same organisation as in previous year.
- Test to see if variables are consistent to reflect same job being worked as previous year.
- Run the tests on the new 'same job' variable as well

Description

- The ASHE form asks 'is the employee in the same job as the previous year'.
 - If 'yes', then ONS copies over information from the previous year (exactly what, we don't know); at least, this is how it was reported many years ago. Now, perhaps not the case – waiting a response from ONS
- In theory, this would allow us to construct longitudinal job links.
- The most recent drop also contains a variable called 'sernol' – 'serial number in last year'. This chain links serial numbers (jobs) across time. This allows us to go better than sjd, and link jobs even if the person has more than one job
- However, we are uncertain where this variable comes from. If created by ONS using some info we don't have (eg name of employer) all good as this is independent of other vars, especially if manually matched; but if it's using a matching algorithm, then it's what we were doing ourselves and we need to identify it as probabilistically matched.

Problem(s):

- Previous work by FR suggest that sjd is not reliable indicator of continued work – significant variations occur in industry, occupation, employer.
- The difficulty is that some individuals might have more than one job – no link in ASHE data between job references in each year, so we need to link ourselves.
- It may well be that ONS no longer uses the 'copy over' strategy, but maybe that it uses it for QA.
- At any event, if it looks like this isn't accurate, we'll need to create our own.
- Problem for multiple job holders: which is the 'same job' being referred to? We can test separately for single job-holders only, then for 'main jobs in multi-job holders, and perhaps also main job irrespective of whether you are single or multi-job holder
- For sernol, the main issue is that it is new (in current ASHE – it did exist many years ago in the NES)

Strategy:

- Load up the years of interest (drop years where eg SIC/SOC change)
- On a rolling two-year basis, for X years (to test whether same problem across time):
 - Only keep those with one job in each year for both years
 - Check to see if sjd is plausible in the sense of whether enterprise, industry, occ etc stays constant
 - Then only keep those with two years of jobs (by job_number)
 - Repeat above exercise
 - Compare sjd and same job as defined by job_number
- Possible next stages (not done yet)

- Repeat, but take main job in multi-job cases, rather than single-job holders
- Repeat for all main jobs (this covers the case where you are multi-job holder in one year and single-job holder in the other)

Findings

First: simple stats on the proportion of those on the same job (either definition) and whether key variables change. Both are taken from the WED dataset (Drop2) as this is the one that can generate job_number.

Table 1 Proportions that have the same feature over two years (single job holders only, same job by ONS definition sjd)

	2009/2010		2013/2014		2017/2018	
	N	Mean	N	Mean	N	Mean
Employer	71,904	96.5%	73,792	97.0%	66,056	96.8%
Home COA	71,904	89.3%	73,792	89.7%	66,056	89.8%
Work COA	71,904	89.4%	73,792	90.2%	66,056	90.2%
Industry	71,904	94.7%	73,792	94.8%	66,056	94.3%
Occupation	71,904	93.3%	73,792	93.1%	66,056	93.4%

Table 2 Proportions that have the same feature over two years (on the same job_number)

	2009/2010		2013/2014		2017/2018	
	N	Mean	N	Mean	N	Mean
Employer	72,312	97.6%	73,257	99.1%	65,220	99.2%
Home COA	72,312	89.7%	73,257	90.2%	65,220	90.4%
Work COA	72,312	90.6%	73,257	91.8%	65,220	92.1%
Industry	72,312	95.6%	73,257	96.5%	65,220	96.2%
Occupation	72,312	94.8%	73,257	94.8%	65,220	95.4%
Same job (ONS defn)	72,312	100.0%	73,257	99.9%	65,220	100.0%

We can break that down a little further by looking to see how many problematic elements there are as some inconsistencies are allowed, using 'employer' as an example:

		In the same job for more than one year?	
		Yes	No
Same employer as last year?	No	Inconsistent	Consistent
	Yes	Consistent	Feasible

For 2017/18 (results for 2009/10 and 2013/14 show similar outcomes), results are as follows. The table plots whether a variable is the same or different, and whether sjd says this is the same job or not. Columns 3-4 show results for only those holding a single job in both years. Column 5-6 show results for those 'same jobs' as defined by job_number; for the latter group sjd is almost always true.

		Single-job holders in both years (73,894 observations)		On the same job in 2 years, as given by job_number (65,218 obs)	
		Same job for more than 1 year?		Same job for more than 1 year?	
		yes	no	yes	no
Employer	Diff.	2.8%	7.8%	0.8%	--
	Same	86.5%	2.8%	99.2%	0.0%
Home COA	Diff.	9.1%	2.9%	9.6%	--
	Same	80.3%	7.7%	90.3%	0.0%
Work COA	Diff.	8.7%	7.8%	7.8%	--
	Same	80.7%	2.8%	92.1%	0.0%
Industry	Diff.	5.1%	6.7%	3.8%	--
	Same	84.3%	3.9%	96.2%	0.0%
Occupn.	Diff.	5.9%	6.3%	4.6%	0.0%
	Same	83.5%	4.3%	95.4%	--

Note: “—” means under 10 observations

There are reasonably reassuring. The inconsistencies in employer feasibly arise from either corporate changes not recognised by the employer filling in the form, or from changes in IDBR referencing. Home COA is more likely to have changed than work COA. Industry and occupation are more likely to have changed whilst still being in the same job, but this may be due to errors in classification by the employer.

Conclusion

These suggest that sjd and job_number can be relied upon to construct links over time (esp the latter). The small inconsistencies could be due to reclassification of industry/occupation/enterprise or (in the case of sjd) because the perception of what that ‘job’ is has changed.

The major concern is that we don’t know where these come from.

If sjd is used to copy over values for the previous year, then testing by comparing one year with the previous is meaningless. The inconsistencies suggest this isn’t the case, so this is positive. Sjd’s value is limited in that it doesn’t distinguish jobs for people in multiple jobs, but as most people only have one job it is a good marker of job continuity.

If sernol is constructed using probabilistic matching using the data available to researchers ASHE, then the findings here are also of limited value as we can carry out the same activity ourselves. If however it is created by ONS using additional information (such as name or other employer details), then again this gives us confidence. Ideally, clerical matching would make this much more reliable, but given the cost (and it is not clear what value ONS would get out of this) this is probably unlikely.