The best of both worlds? Linking survey and administrative data for labour market research

Dr Darja Reuschke
City-Region Economic Development Institute
Birmingham Business School



UNIVERSITY OF BIRMINGHAM

UKRI Economic and Social Research Council

ADR UK Data-driven change

Content

1. The Annual Survey of Hours and Earnings (ASHE)
2. Data access
3. The linked datasets: SA and PAYE
4. My project: multiple employment
5. Opportunities and limitations

The data

- **ASHE-GB (excludes Northern Ireland)**
  - Sample of 1% of employee jobs in GB (unit of analysis: jobs)
  - Information is provided by employer for reference day in April
  - Employer non-response, not at random
  - Main sample: PAYE jobs based on National Insurance Number (NINo)
  - Inclusion in the sample is based on the 7th and 8th digit in the NINo being 14.
  - NI numbers remain in ASHE sample -> longitudinal element
  - Multiple (PAYE) jobs of the same NI number are sampled

Data access

- Office for National Statistics; His Majesty's Revenue and Customs, released 01 August 2024, ONS SRS Metadata Catalogue, dataset, **Annual Survey of Hours and Earnings linked to PAYE and Self-Assessment data - GB**, https://doi.org/10.57906/566k-5q15
- Secure Research Service of the ONS
- Application for an accredited project using PASS https://ons.metadata.works/browser/landing
- Longitudinal ASHE file (1997-2022)
- ASHE-SA datafile
- ASHE-PAYE datafile
- Link ID: *Piden*
- No link with other SRS datasets including Census data

Linked datasets: SA

ASHE-SA 2011-2017

- Self-employment income and other (taxable) income
- Aggregate for tax years (2010/11 – 2016/17)
- From the Self-Assessment (SA), all returns that match the sample pool for ASHE have been included in the dataset – i.e. all returns with a NINo with 7th and 8th digits being 14. That is **798,427 records**.
- SA_subset_master file contains n= 806,802 observations.
- From this dataset, **n=119,151** can be linked with the ASHE longitudinal file = match rate of 8.45%

ASHE-SA linked data by ASHE year

ASHE year	obs in ASHE	linked with SA	match rate
2010	171,861	13,096	7.62
2011	180,009	13,652	7.58
2012	172,905	14,346	8.30
2013	175,518	15,295	8.71
2014	181,041	16,170	8.93
2015	178,980	15,877	8.87
2016	174,903	15,530	8.88
2017	174,781	15,185	8.69
N obs (total)	1,409,998	119,151	8.45
N individuals	330,455	45,826	

- Sample of ASHE individuals with profit from self-employment: **n=26,006** (1.84%)
- % of individuals with 2nd jobs as self-employed >£1,000pa is higher in UKHLS

UNIVERSITY OF BIRMINGHAM

How often do ASHE individuals have linked SA records?

counts (year)	with SA data	with s/emp income
0	284629 (86.1%)	316287 (95.7%)
1	19,954	9,198
2	8,458	2,157
3	5,376	1,060
4	3,899	634
5	2,832	450
6	2,213	302
7	1,633	220
8	1,461	147
	330,455	330,455

Linked datasets: PAYE

HM Revenue and Customs Pay As You Earn (HMRC PAYE)

ASHE-PAYE 2014-2018

- RTI from employer to HMRC
- Monthly and weekly dataset
- 18M payslips – with no other information than pay; no hours or any information on the job; includes pension
- Becomes meaningful through link with ASHE (gender, age)
- 690,939 ASHE individuals aged 16+
- BUT 21% of ASHE individuals aged 16-64 do not have payslips in PAYE data

My project

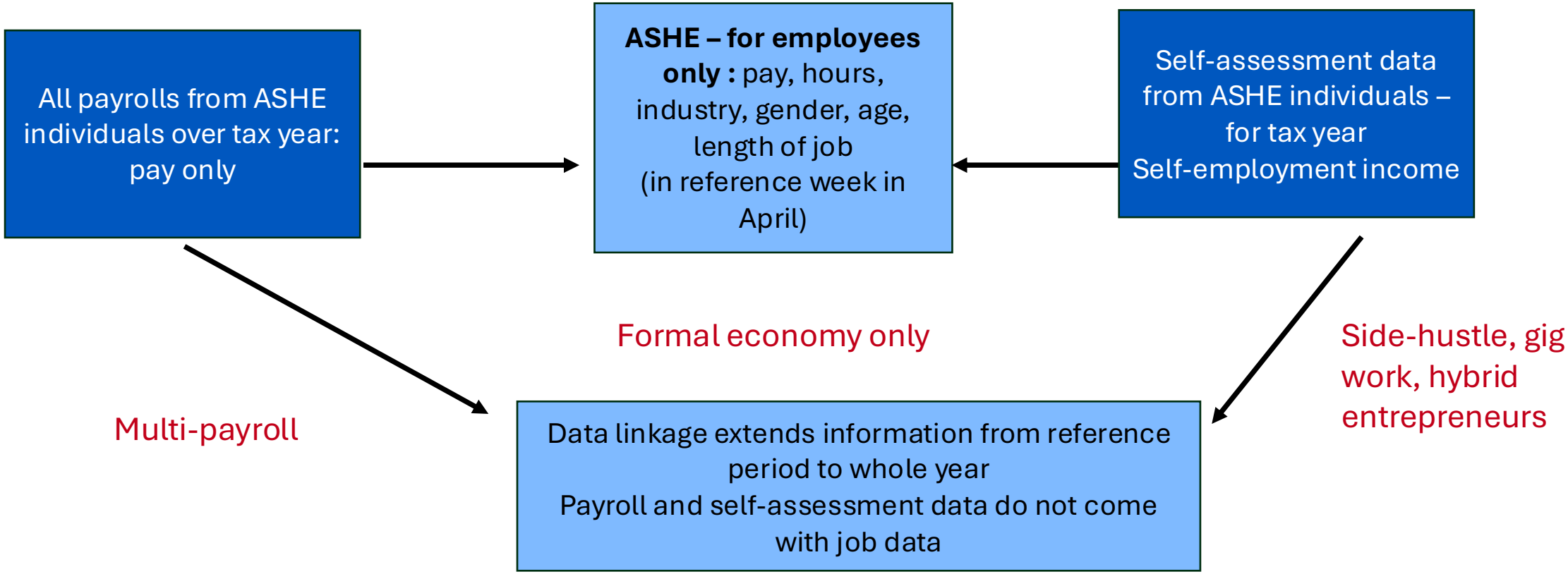
- Multiple jobs as an employee in official labour market data low:
  - Labour Force Surveys & Annual Survey of Hours and Earnings: ca. 3% (women: 4%)
  - EU Workforce: 4% (Eurofound, 2020)
- Higher in non-governmental survey data
  - UK Understanding Society: 9% women and 6% men (2023-24)
- Reference week/day - problematic for ‘nonstandard’ employment
- Who reports ‘second’ jobs in social surveys? What do people consider or report as ‘second’ job?

UNIVERSITY OF BIRMINGHAM
UKRI Economic and Social Research Council
ADR UK Data-driven change

Objectives

- To use data not limited to a reference period and not affected by response bias
- To study influencing factors on multiple employment among women vs men
  - by multiple employee jobs vs self-employment activities
  - Regional labour market conditions
- Research questions
  - Financial constraints?
  - Job insecurity?
  - Labour market demand

Data for my project



Linked administrative data

- PAYE data from 6 April 2014 to 5 April 2019
- Real time payroll data (weekly)
  - Identification of individuals with two or more payrolls from different employers on a weekly basis
  - Measure what people do (within limits), not what they report as 2nd job
  - Measure of paid employment activity but not 'jobs'
- Identification of multiple jobbing (multi-payroll)
  - Duplicate payslips of same individual in same year-month
  - Excluding pensions & payslips with negative or zero value
- Number of months per year with multiple payslips (0-12)
- Self-employment 'side' income as profit from self-employment
- But only 2014-2016 ASHE years with both SA and PAYE linkage

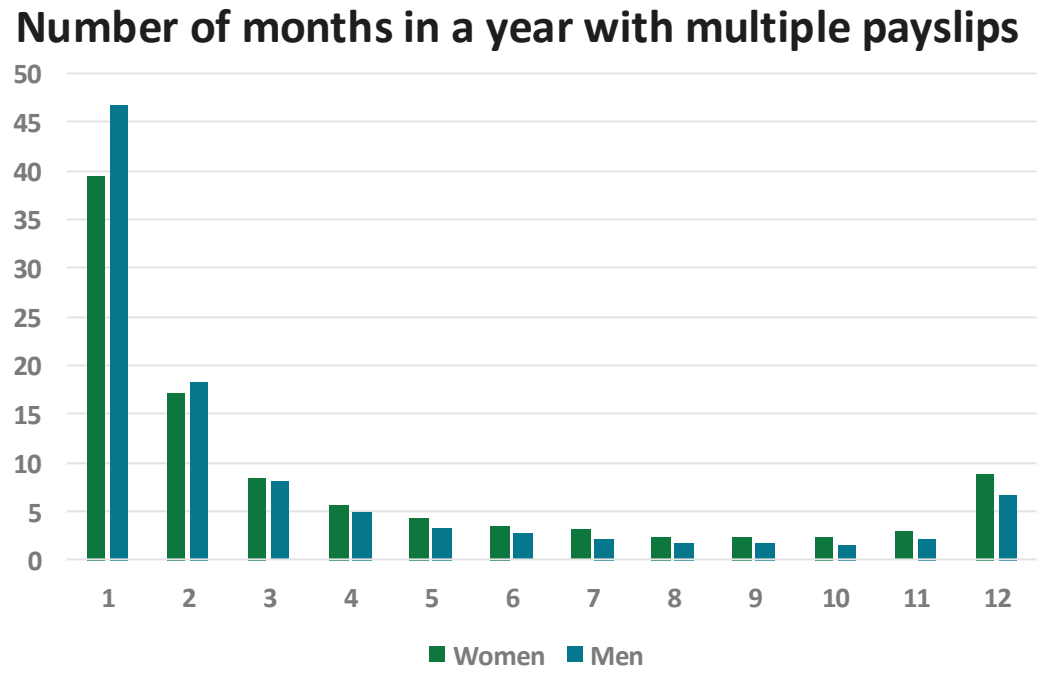
UNIVERSITY OF BIRMINGHAM
UKRI Economic and Social Research Council
ADR UK Data-driven change

Multiple jobs discovered in data

Classifications	Women	Men
2+ jobs in ASHE	3%	1.4%
2+ payslips in tax year but 1 job in ASHE	15.5%	13.8%
1 job in ASHE and 1 payslip throughout the tax year	81.5%	84.8%

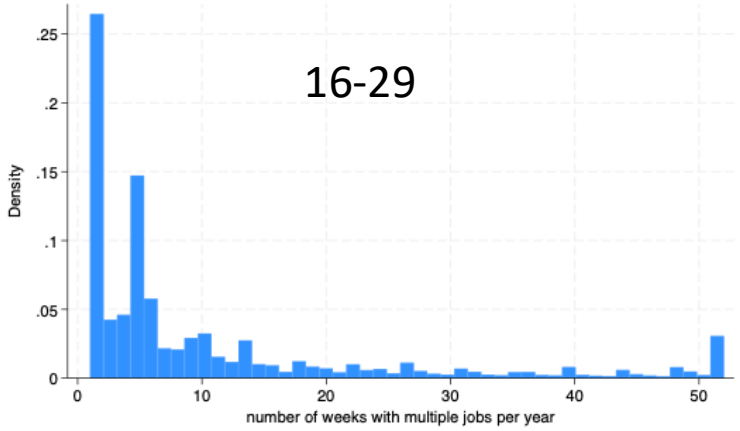
Note: Unweighted data.

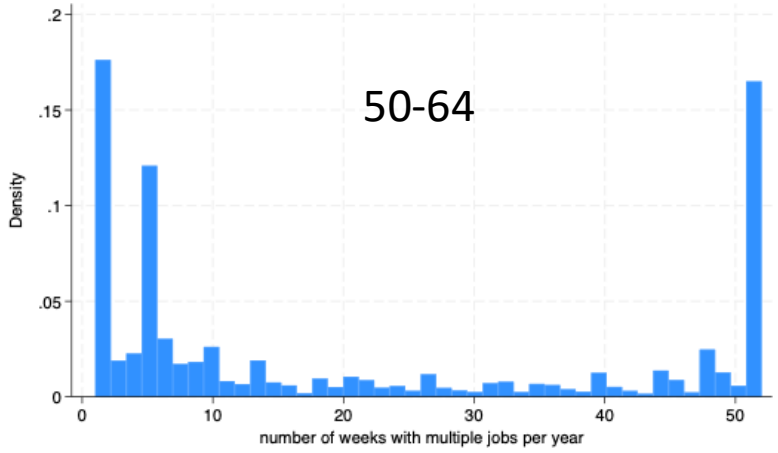
- ➔ Multiple employment in ASHE is low.
- ➔ In real-time payslip data, multiple employment (defined as having more than one payslip in a month in a year) is much higher. This form of multiple employment is not captured ('invisible') in social survey data.
- ➔ Many only have more than one payslip in one month only.

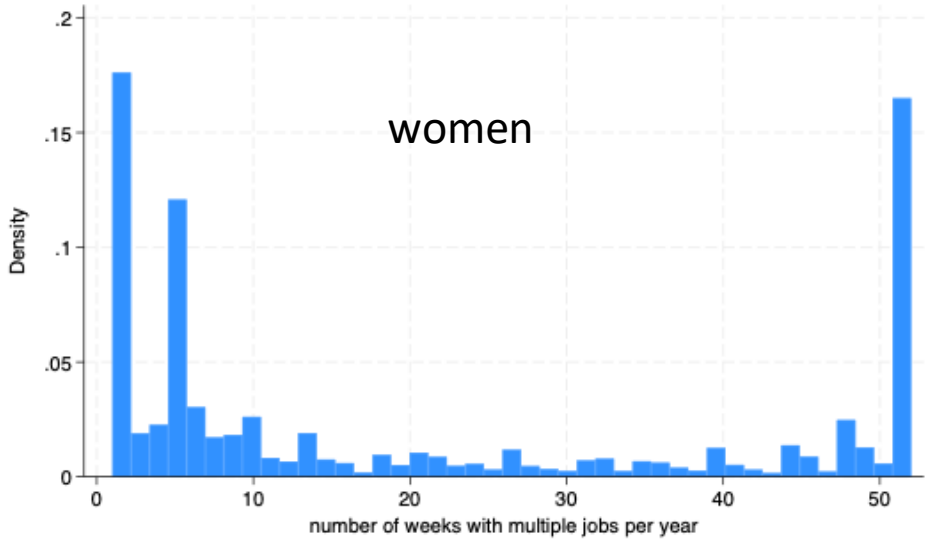


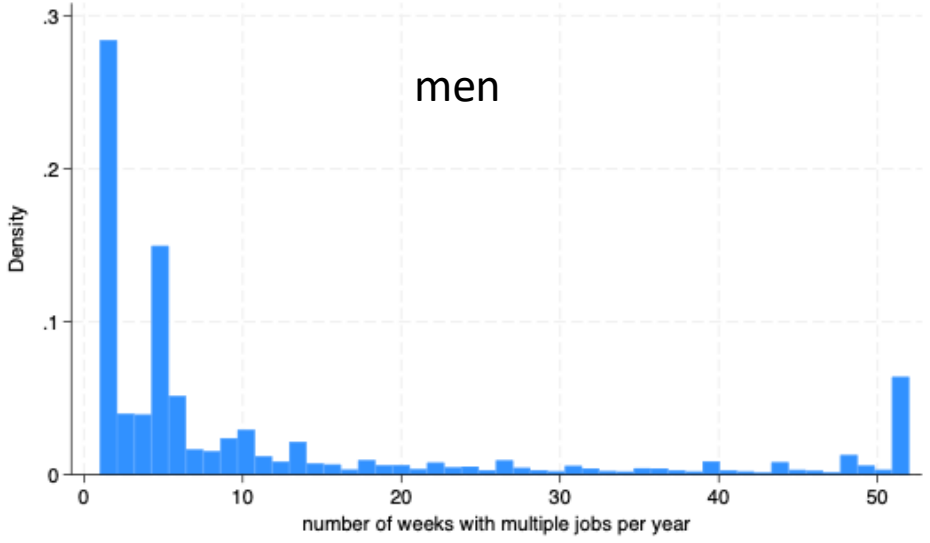
UNIVERSITY OF BIRMINGHAM

Number of weeks



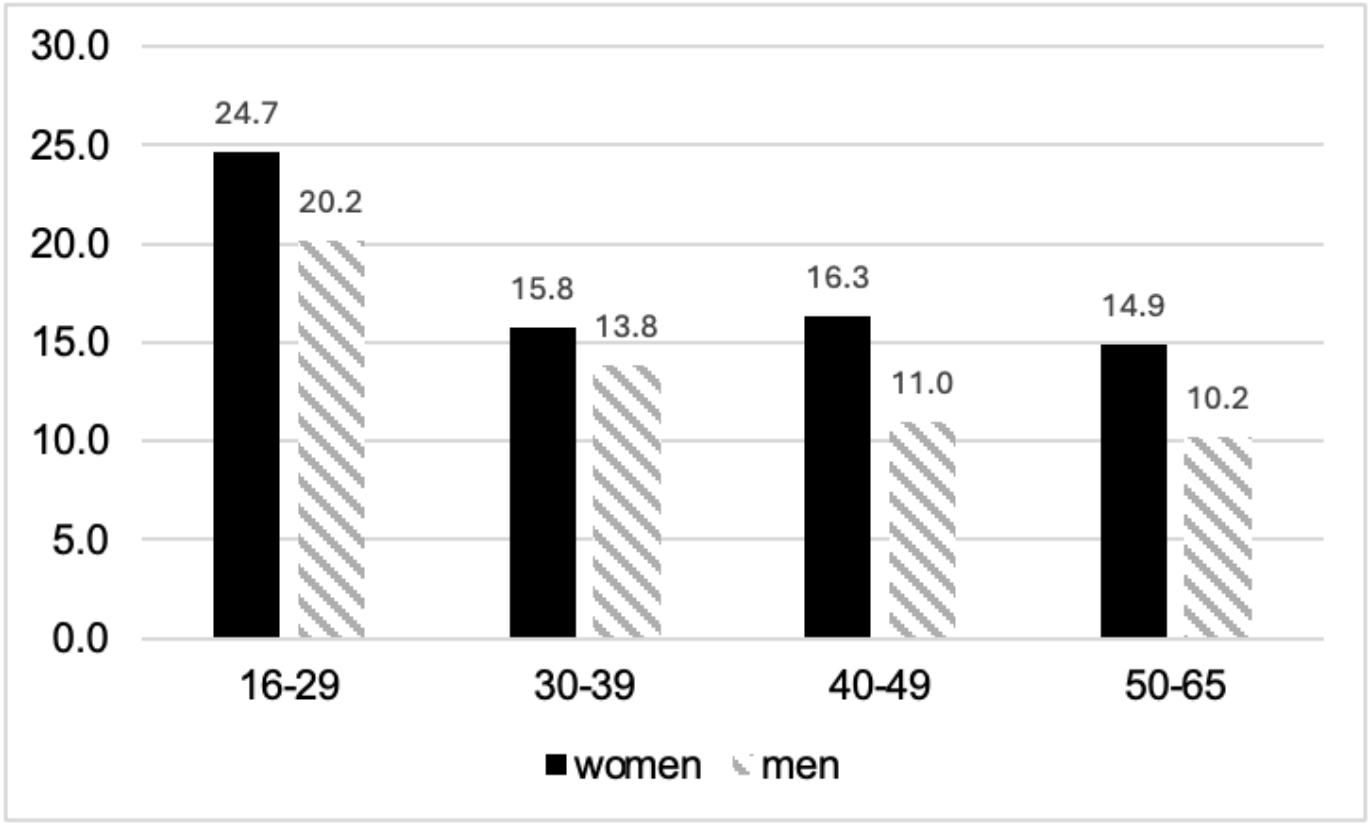






UNIVERSITY OF BIRMINGHAM

High level of multiple jobholding with LFS weights



Multiple jobholding in the linked ASHE-HMRC PAYE data, by gender and age groups, 2018-19, in percent, LFS population weights

Comparison: Estimates of second earnings of households in the USA: 18% (Scott et al., 2020)

UNIVERSITY OF BIRMINGHAM

UKRI Economic and Social Research Council

ADR UK Data-driven change

Analysis samples

Women

Employee & self-employment income	multi-payroll (>5 weeks/year)
n=360,624	n=356,566
1.9%	12.2%

Men

Employee & self-employment income	multi-payroll (>5 weeks/year)
n=328,960	n=326,845
2.1%	7.9%

Conclusion: Opportunities and limitations

- A lot of geographical identifiers:
  - Small geographies for both workplace and residence
  - Linkage of geographical data to model labour demand: regional/local unemployment, GDP, median wage
  - BUT: workplace is prefilled by ONS
- Longitudinal nature of ASHE
  - Investigation of transitions and progression
  - BUT large nonresponse with no longitudinal weights
- Job data only in ASHE
- Data quality:
  - Why do one-fifth of ASHE individuals not have PAYE data?
- Continuation of data linkage? Linkage with Census (ethnicity)?

Thank you!

Disclaimer: “This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners.”

Office for National Statistics; His Majesty's Revenue and Customs, released 01 August 2024, ONS SRS Metadata Catalogue, dataset, **Annual Survey of Hours and Earnings linked to PAYE and Self-Assessment data - GB**, https://doi.org/10.57906/566k-5q15

Acknowledgement: The research was funded by an ADR UK fellowship (ES/Z503149/1) and conducted in collaboration with the UK WBG and Prof Tracey Warren (University of Nottingham).

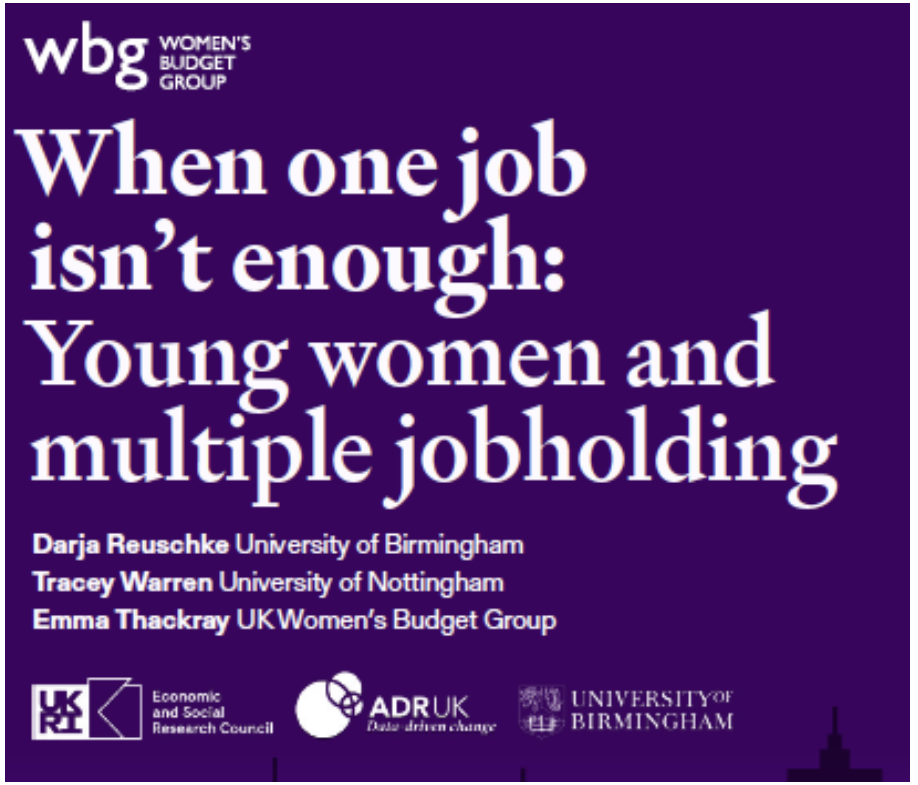
UNIVERSITY OF BIRMINGHAM

UKRI Economic and Social Research Council

ADR UK Data-driven change

Darja Reuschke: d.reuschke@bham.ac.uk

Report: https://www.wbg.org.uk/wp-content/uploads/2026/04/WBG-Young-Women-Report-Digital.pdf



Data Insight: Linked administrative data reveals much higher level of multiple jobholding in Britain than previously thought

https://doi.org/10.5281/zenodo.23571668

Multiple Job holding: A critical quantitative methodology
https://futuresofwork.co.uk/2026/07/08/multiple-job-holding-a-critical-quantitative-methodology/

UNIVERSITY OF BIRMINGHAM

Economic and Social Research Council

ADR UK Data-driven change