Young People’s Earnings Progression and Geographic Mobility

Bonnie Lewis & Tom Odell
Policy Evidence and Analysis Team, ONS
Email: policy.evidence.analysis@ons.gov.uk
Article available here

05 June 2019
Key messages

1) Median earnings growth between 2012 and 2016

83.4%
16-29 year olds

5.6%
Working age population

2) Young people’s earnings progression is related to geographic mobility. The highest average annual growth in earnings was 22% for those moving to London between 2012 and 2016

3) Compared to men, young women aged 16 to 29 years had lower levels and growth of annual earnings regardless of qualifications between 2012-2016

Link to article – published October 2018
1) Background & dataset information
2) Earnings progression measures
3) Movers analysis
4) Regression modelling
5) Next steps
Project background

Disclaimer: These Research Outputs are experimental statistics on earnings progression. This analysis is based on a feasibility study exploring the use of administrative data linked to the 2011 Census.
What? – Using a linked dataset

Census 2011

Individual information: Age, sex, ethnicity, country of birth, employment status, occupation, industry, health and disability status.

Household Information: Household structure – age of youngest child, location (LA), family relationships in household

Customer Information System
Location of home address (LAD)

National Benefit Database
Information for 13 benefits

Pay As You Earn
Income received for 2011

Tax Credits

Universal Credit

Personal Independence Payment

Single Housing Benefit Extract

Child Benefit
• **Who isn't included** – Self employed individuals; new migrants to the UK (post-2011).

• **Geography** – This analysis is restricted to those resident in England and Wales at the time of Census 2011. Local Authority geography is also taken from the CIS to analyse impact of moving on earnings progression.

• **Earnings** – As this analysis only considers annual earnings, growth may be due to increased hourly pay, or an increase in number of hours worked.

• **What is meant by progression** – This analysis only covers earnings progression. Changes in occupation or improvement to conditions of work which may be considered progression are not analysed here as this information is not currently available on the dataset.

• **Occupation and Education** – taken from Census 2011. Individuals may have gained additional qualifications and changed occupations during the period of analysis which will likely impact earnings.

### Population

- **Population** – working age population (16 to 64 years) for each year, 2011/12 to 2015/16. This is a longitudinal dataset that matches individuals over time. Individuals are contained in the dataset if they were present on the 2011 Census and are on PAYE. The dataset population increases year on year by a similar margin to the ASHE number of jobs (which was approximately 24m in 2011 and 26m in 2015).
### What? – Using a linked dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Benefits</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| Annual Survey of Hours and Earnings | ONS’s most comprehensive survey on earnings and covers the earnings of paid employees in the UK | - Longitudinal weighting element is currently only available over two years. This is ineffective when considering lifetime trends.  
- The ASHE survey is a business survey and so DOES NOT include a range of personal characteristics that have been identified in prominent literature as vital to progression such as education or parental status. |
| Labour Force Survey/Annual Population survey |  
- The LFS holds very good information on education and training which has been identified as a key driver of in work progression.  
- It also holds considerable information on personal characteristics |  
- Self reported earnings  
- The longitudinal element only covers 5 quarters which is often not enough time to look at significant barriers/enablers to progression. |
| Longitudinal Census              | A very large variety of personal characteristics covered                  |  
- Does not have information on income, so progression is defined via occupation and hours worked  
- Over a 10 year time frame so we don’t know what they are up to in between. Also 2001 and 2011 so the analysis maybe considered out of date? |
| Census Link dataset             |  
- Large sample size therefore the results are likely to reflect the population as a whole  
- 5 year time frame, how income has changed for individuals and households over this time  
- Both household and individual level so that family circumstances are taken into account  
- Define labour market progression as increase in earnings and reduction in benefit/tax credit claims.  
- Can also be expanded for the full employment work and movements into and out of the labour market rather than just in work progression |  
- We do not have information on their current occupation and employment status  
- No information on hourly earnings |
Overview of workplan

January 2018
- Deconstruct dataset for analysis
- Conduct sense checking & descriptive analysis
- Analysis of different progression measures
- Investigate geographic mobility and relationship to pay
- Conduct regression analysis

October 2018
- Write article and deliver analysis to DWP & RDU
Relative earnings growth
(Social Mobility)
Relative earnings growth methodology

- Social mobility measure

- Earnings deciles of working age population

Progressed (in employment)
Those who had PAYE data greater than zero in both years and had moved up at least two deciles since the previous year

Progressed (into employment)
Those who had PAYE data in the specified year but did not have PAYE data in year one

Not progressed (in employment)
Those who had PAYE data greater than zero in both years, but had not moved up at least two deciles since the previous year

Not progressed (out of employment)
Those had no PAYE data in the specified year, but did in the previous year
Relative earnings growth findings

Between 78% and 81% of working age adults did not experience year-on-year pay progression of at least two earnings deciles.

Younger people were also more likely to experience pay progression at the start of their career, with 19% of those aged 16-24 progressing by two or more deciles between 2015 and 2016.
Annual earnings growth
The earnings growth rate is calculated for employees with PAYE earnings greater than zero in two consecutive tax years. Growth in earnings between the two tax years is calculated for each individual using the formula:

\[
\text{Earnings growth} = \frac{\text{Year 2 earnings} - \text{Year 1 earnings}}{\text{Year 1 earnings}}
\]

The median of all growth rates is then taken, to provide an estimate of “typical” earnings growth for employees between the two consecutive tax years. Earnings growth is adjusted for the effects of inflation using the Consumer Prices Index including owner occupiers’ housing costs (CPIH).

Methodology differs to that normally produced using data from the Annual Survey of Hours and Earnings (ASHE). ASHE calculates the median earnings for a population in two consecutive tax years, then calculates the growth between these median levels of earnings.
Young employees have seen the greatest earnings growth between 2012 and 2016.

Among other factors, this may also have been influenced by the rise in National Minimum Wage for this group each year during this period.

Main driver of growth for young people would likely be changing hours & moving from casual jobs to more permanent careers.
Relative low annual pay
(Resolution Foundation)
**Escapers**
Those who earned above the relative low annual pay threshold in every year from TYE 2014 to TYE 2016 suggesting they have progressed onto higher wages

**24.1%**

**Cyclers**
Those who fall between the first two categories, moving above the threshold for some but not all years between TYE 2013 and TYE 2016

**42.2%**

**Stuck**
Those who were below the relative low annual pay threshold in every year between TYE 2013 to TYE 2016

**21.2%**

**Exited**
Those who are no longer present in the data at the end of the period and their progression is unknown

**12.5%**

Based on Resolution Foundation definitions. Figures for 2.2 million 16-24 year old's 2012-2016.
Relative low annual pay methodology & findings

Relative low annual pay status by age group

16-24 year old's were most likely to be cycling in low pay.

People aged 36 to 55 were more likely to be stuck in low pay.

56 to 64 year old's tended to exit the dataset (likely due to retirement)
Who moves? – Where do they move? – Does this impact on earnings growth?
Who moves local authority?

Young employees are more likely to be moving to London and surrounding LA’s, while older people (46-64) tend to move to coastal and rural regions.
Who moves local authority?

Meanwhile, those who are highly skilled are also more likely to move.
CHARTS SHOW RAW NUMBER OF MOVERS TO CITY REGIONS BETWEEN 2012 AND 2016 (E.G. FROM CORNWALL).

OUTPUTS ROUGHLY MATCHED PUBLISHED ONS INTERNAL MIGRATION FIGURES – DIFFERENT ACCOUNTED FOR IN SOURCES USED.

MAIN FINDING: YOUNG PEOPLE TEND TO MOVE TO CITY REGIONS CLOSEST TO THEM.
Moving is associated with earnings growth

Young people aged 18-29 experience the greatest earnings growth if they move to London.

For those who do not move, the greatest earnings growth is experienced by those in London and the South East.

We conducted analysis of people moving to city regions across England & Wales...
Moving is associated with earnings growth

Highest growth tends to be experienced for young people moving from more distance areas – particularly movers to London.

In some cases, earnings growth is lower once young people have moved into a city region (LA’s surrounding Cambridgeshire & Peterborough).
Disparities analysis for non-movers
The effect of socio-demographic disparities for non-movers

This graphic demonstrates how it was possible to visualise the data in one interactive.

These young people were non-movers and had been employed consistently between 2012 and 2016.

Each line represents a combination of socio-demographic characteristics (e.g. Asian, male, school educated, aged 26 to 29, living in the east)

Characteristics could be selected to filter the lines, and a further characteristic can be selected to compare between the lines.
The effect of socio-demographic disparities for non-movers

For young employees (16-29), BAME men typically experienced some of the highest earnings growth between 2012 and 2016, but had low initial earnings.

White women were most likely to have poor earnings growth but had reasonably average initial earnings.

For those aged 22-29, degree educated white men in London had the highest median earnings at the start and end of the period.
Regression modelling

Research questions
1) What factors are associated with escaping low pay?
2) What factors are associated with moving local authority?

How?
Sampling – stratified by geographic region to account for pay differences and increased the size from 1,000 to 10,000 to ensure representative sample/not using whole population.
Populations in the final sample of 10,000 subjected to sense checks.

Multicollinearity - correlation covariance matrix identified correlating variables (above 0.7). Variables removed which were more interpretable/important in the model (e.g. disability included over health)

SAS – using backwards stepwise modelling, where variables with the highest p-values were removed gradually until only variables significant at the 0.05 level remained.

Forward stepwise method was also tested and produced a similar model but introduced suppressor effects.

Independent variables taken from 2011; dependent variables based on outcome by 2016.
Regression modelling

Testing the models

_Goodness of fit_ - Likelihood Ratio Chi-Square Test, the Score Chi-Square Test and the Wald Chi-Square Test were all significant at the less than 0.001 level, meaning that at least one of the regression coefficients in the model was not equal to zero.

_R-squared values_ – testing how well the logistic regression equation predicts the outcome with and without the independent variables.

_Accuracy rates_ - tells us what percentage of responses a model predicts correctly. A percentage of 50% would mean our model is no better than a random guess.

_Hosmer and Lemenshow test_ - checks the null hypothesis that the model fits the data well. This test is highly sensitive, and with different samples we get different results for the same model. The greater number of categories a model has, the more power there will be to reject the fit.
### Regression modelling – Model fit

<table>
<thead>
<tr>
<th>Model 1 – Factors associated with escaping low pay</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared value</td>
<td>13%</td>
</tr>
<tr>
<td>Accuracy rate</td>
<td>70%</td>
</tr>
<tr>
<td>Hosmer and Lemenshow 'goodness of fit' test</td>
<td>0.5377</td>
</tr>
</tbody>
</table>

**Escaping low pay:**
“Whether an individual on low pay in TYE 2012 moved off low pay for consecutive years 2013-2016”

<table>
<thead>
<tr>
<th>Model 2 – Factors associated with moving local authority</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared value</td>
<td>11%</td>
</tr>
<tr>
<td>Accuracy rate</td>
<td>71%</td>
</tr>
<tr>
<td>Hosmer and Lemenshow 'goodness of fit' test</td>
<td>0.0022</td>
</tr>
</tbody>
</table>

**Moving local authority:**
“Whether an individual had moved between local authorities at least once between 2012 and 2016”
Regression modelling

Interpretation of results

We used the odds ratio, interpreted as: for a one-unit change in the predictor variable, the odds ratio for a positive outcome is expected to change by the respective coefficient, given the other variables in the model are held constant.

Reference categories were set at a theoretically relevant response.

The interpretation differs for the household earnings variable as it was the log of this variable that was included in the model. Instead of a one-unit change in this variable, the value of the odds ratio is associated with a 10% increase in household earnings.
## Model 1 – Factors associated with escaping low pay

### Odds of escaping low pay

<table>
<thead>
<tr>
<th>Status</th>
<th>Reference</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disability</td>
<td>Not limited</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Limited a little</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Limited a lot</td>
<td>2.5</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>White</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Other ethnicity</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Black ethnicity</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Asian ethnicity</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Mixed ethnicity</td>
<td>0.5</td>
</tr>
<tr>
<td>Skill level</td>
<td>Low skill</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Upper middle skill</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Lower middle skill</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.8</td>
</tr>
<tr>
<td>Industry</td>
<td>Other service industries</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Public admin, education and health</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Banking and finance</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Transport and communication</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>Distribution, hotels and restaurants</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Construction</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Manufacturing</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Energy and water</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Agriculture, forestry and fishing</td>
<td>2.5</td>
</tr>
<tr>
<td>Household type</td>
<td>Single without children</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Couple with children</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Single with children</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Couple without children</td>
<td>2.5</td>
</tr>
<tr>
<td>Benefit status</td>
<td>No benefits</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Working age and other benefits</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Other benefits</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Working age benefits</td>
<td>2.0</td>
</tr>
<tr>
<td>Age</td>
<td>Not moved</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Moved local authority</td>
<td>0.8</td>
</tr>
<tr>
<td>Qualification</td>
<td>No qualifications</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Level 4 qualifications</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Level 3 qualifications</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Level 2 qualifications</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>Level 1 qualifications</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>Apprenticeships</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>Other qualifications</td>
<td>6.0</td>
</tr>
<tr>
<td>Reference: No qualifications</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

- Those of black ethnicity are 60% more likely to escape low pay compared to people of white ethnicity (1.6 times higher).
- Region was not statistically significant in the model.
- However, moving local authority increases the likelihood of escaping by 20% (0.8 times lower for those who do not move).
- Having level 4 qualifications means you are twice as likely to escape low pay compared to having no qualifications (2.1 times higher).
Model 2 – Factors associated with moving local authority

Londoners are significantly more likely to move than people from any other region in England and Wales. People in the South East are 28% less likely to move than those from London; people from the North East are 60% less likely to move.

Women are 15% less likely to move than men.

Having level 4 qualifications means you are twice as likely to move local authority compared to having no qualifications.
Impact
Women’s private pension pots outstrip men’s, says ONS

By Michael Klimes | 23rd October 2018 2:44 pm

The gap between what women earn and save compared to men continues to grow according to Office for National Statistics figures released today.

Analysis for young people aged 16 to 29 shows the level and growth of median annual earnings were generally lower for women than men whether they were degree or school-educated or had no qualifications.

Men in this age group with a university education in London earn, on average £27,081, while women with same background earn about £24,767.

Historically, men have had a substantially higher median pension value than women.

However, that trend has begun to reverse in recent years.

Numbers from the ONS show women aged 16 to 24 have about £3,900 in a pension compared to men who have about £3,000.

The gap widens at ages 25 to 34 when women have about £14,400 in a
Next steps
What did we learn and what’s next?

Related work:
- Ethnicity Pay Gaps – publishing in the summer
- Disability Pay Gaps
- Economic Inequalities

New dataset & analysis:
- Admin Data Research UK (Contact: adrcuration@ons.gov.uk)
  - Creating evidence for government policy by joining up and utilising administrative data already created by government and public bodies across the UK
  - BIDs may be available via SRS – long term/2020
- Potential to link Census + BIDS + HESA -> filling the evidence gaps left by this work
  - What would be of interest using this dataset?
    - Progression & productivity?
    - Progression & changes to level of education
  - Linkage is a long process, but research questions would be welcome in the meantime!
Thank you for listening!
Any questions?