Office for National Statistics

Compendium

Overeducation and hourly wages in the UK labour market; 2006 to 2017

This article examines overeducation in the UK labour market using Annual Population Survey (APS), for 2006 to 2017 including analysis on the relationship between overeducation and wages.

Contact:
Maja Savic
economic.advice@ons.gov.uk
+44 (0) 207 592 8698

Release date:
29 April 2019

Next release:
18 July 2019

Table of contents

1 . Main points

- In 2017, around 16% of all those in employment aged 16 to 64 years were overeducated (had more education than required for their job); the corresponding figure for graduates (with first degree or equivalent) was around 31%.

- In 2017, 21.7% of those who graduated before 1992 were overeducated, whereas the corresponding figure for those who graduated in 2007 or later was 34.2%.

- There is a wage penalty associated with overeducation, although overeducated employees earn positive return on wages, this is significantly lower compared with those who are matched to their jobs.

- In 2017, the overeducation rate was similar for women and for men, however the wage penalty for overeducation was somewhat higher for men than for women; this suggests that overeducation does not contribute to gender pay gap.

- Recent graduates experience lower pay penalty on overeducation compared with non-recent graduates.

2 . Introduction

This article examines overeducation in the UK labour market using Annual Population Survey (APS), for 2006 to 2017. Overeducation is a form of mismatch where a person can be overeducated if they possess more education than required for the job. Overeducation is a form of resource underutilisation, which may have implications for the individual, firm and the economy. It can also be seen as a form of underemployment, hence contributing to the extent of labour market slack.

The existence of overeducation has been explained by drawing emphasis on the role of human capital. This means that workers may substitute education for the lack of previous work experience, accepting jobs requiring less education than they possess, until they acquire further experience.

This explanation is consistent with the career mobility argument, whereby individuals may forgo higher wages in the early stage of their careers, only to experience upward income mobility later in their careers. Both the human capital and career mobility explanations perceive overeducation as a temporary situation, which is likely to cease once the worker has gained sufficient experience.

An alternative explanation for the existence of overeducation, put forward by the dynamic labour market view, relates to the presence of information asymmetries which may cause search frictions. In other words, and according to dynamic labour markets view there will always be some degree of mismatch between education achievement and job requirements due to incomplete information which prevents efficient matches between educated workers and employees. Some peoples' education will exceed requirements and they will be overeducated (underemployed), others will be in the opposite situation.

Overeducation may also arise in the context of labour market distortions if the supply of graduates exceeds the demand and crowds out job opportunities for the less educated. Employers may prefer hiring graduates as this is likely to reduce training costs, while less educated workers may become unemployed. In this case overeducation could be a more long-term phenomenon.

At the individual level the relationship between overeducation and productivity is often captured through its impact on wages. Existing evidence shows that overeducated workers are likely to earn lower returns relative to similarly educated individuals whose jobs match their education (Hartog, 2000; Green and Henseke, 2016). At the firm level, there is some evidence to suggest that overeducation is associated with lower productivity (McGuinness, 2006; Green, 2016). McGowan and Andrews (2015) find a similar negative relationship at the industry level, for a large sample of Organisation for Economic Co-operation and Development (OECD) countries.

Our article starts with a detailed evaluation of the extent of overeducation in the UK, accounting for age, gender, and regional differences. We also present a detailed analysis by type of first degree and status of younger graduates. The cost of investing in tertiary education and uncertainties related to future job opportunities, make this extension particularly relevant.

Our study aims to answer the following research questions:

- what is the incidence of overeducation in the UK labour market by sex, age and region?
- what is the incidence and persistence of overeducation for graduates (with first degree or equivalent) by type of degree subject?
- what is the relationship between overeducation and wages? Do results for women and men differ?
- are younger (recent) overeducated graduates earning lower wages, compared to older (non-recent) overeducated graduates?

Two main assumptions underlie our analysis. First, we assume that education is a measure of workers' abilities, hence we do not directly distinguish between overeducation and over-skilling. This distinction can be important if highly educated workers are employed in low-level jobs because they only possess low skills, despite their level of education.

Second, we take wages as our measure of productivity, hence this article does not analyse the impact of overeducation on productivity measured in terms of output per worker or hour worked or total factor productivity.

In this article, using the Annual Population Survey (APS), we broadly follow a statistical method used by the International Labour Organisation (ILO) to compare the distribution of educational attainment of those in employment in the UK against the average educational attainment level for their occupation.

Each individual is assigned a status based on whether their own level of education falls within or outside of the range for their particular occupation and age group. The range is defined as being one standard deviation above and below the mean level of educational attainment.

This method, will, by construction, always result in a proportion of workers who can be classified as:

- matched (individuals whose highest qualification falls within one standard deviation of the average level of educational attainment for their occupation)
- overeducated (individuals whose highest qualification is above one standard deviation of the average level of educational attainment for their occupation)
- undereducated (individuals whose highest qualification is below one standard deviation of the average level of educational attainment for their occupation)

Aggregating these groups over all occupations gives an estimated matched, over and undereducated rate for the whole economy.

Notes for Introduction:

1. A different type of mismatch, which we shall not pursue in our analysis, is that of “horizontal mismatch” where the subject or type of education does not match the job requirements.

2. We should emphasise that the statistical method, by its construction, permits the average educational attainment to increase across all occupations if participation in education and the average level of educational attainment in the population increases. The effect on the degree of matching across the whole economy is therefore dependent on the age composition of each occupational group and the distribution of older and younger workers across occupations. So, to mitigate a potential age composition bias, we construct estimates of average educational attainment by three-level Standard Occupational Classification (SOC) group and by two age groups: 16 to 35 years and 36 to 64 years. We present our methodology for measuring education mismatch and overeducation in Appendix 1.

3 . Results-descriptive analysis

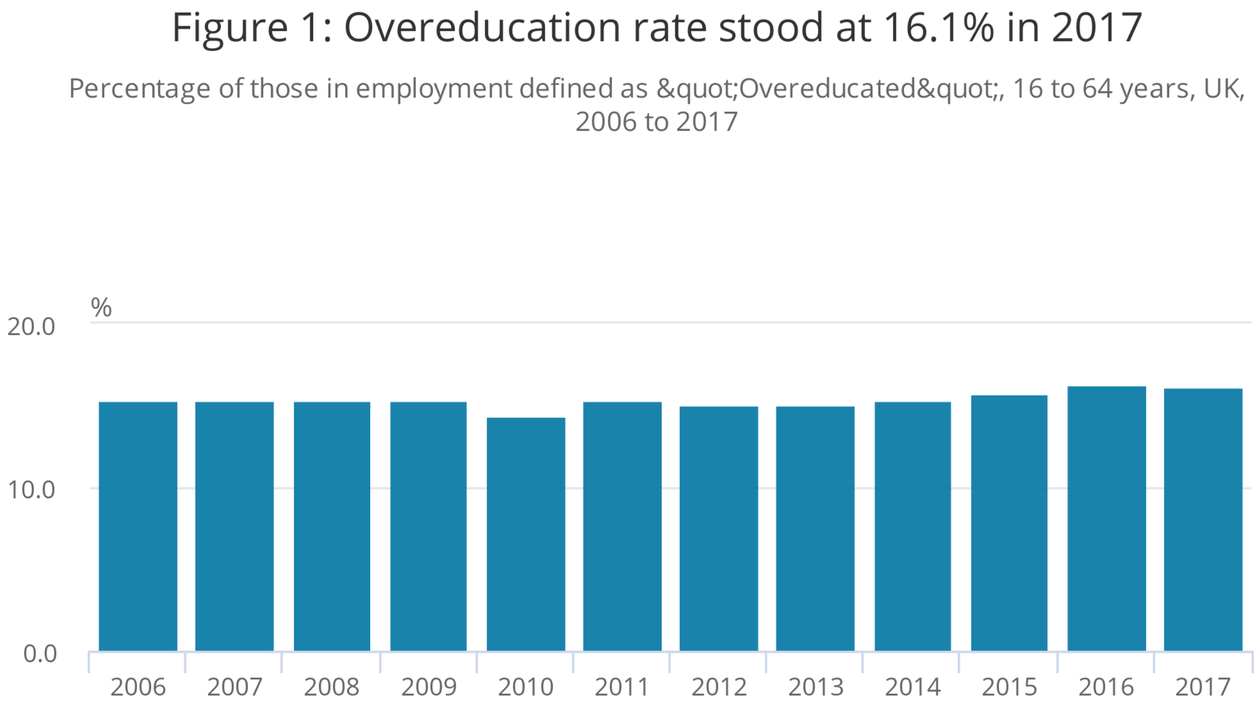
All in employment aged 16 to 64 years

Aggregate results show that in 2017, 68.4% of those in employment had a level of education close to the average of their job. Figure 1 presents percentage of those in employment who are classed as overeducated. From 2013, the overeducation rate began to rise reaching a peak of 16.3% in 2016.

This may be the result of the increase in the number of individuals attaining a first degree or equivalent. In fact, between 2013 and 2017 the total number of graduates in employment increased by 14.8%. The increase in overeducation could also mirror an increasing competition for higher skilled jobs and a surplus of candidates. From 2016 to 2017 the overeducation rate decreased slightly, by less than 0.2 percentage points, and in 2017 it stood at 16.1%. Over time, competition amongst graduates may rise the level of education required for the job, hence increasing the threshold. This may lead to a decrease in the number of overeducated employees.

Figure 1: Overeducation rate stood at 16.1% in 2017

Percentage of those in employment defined as "Overeducated", 16 to 64 years, UK, 2006 to 2017



Source: Annual Population Survey – Office for National Statistics

Notes:

1. The data for estimates prior to 2011 were collected on the previous SOC basis (SOC 2000) and have been mapped to an equivalent SOC 2010 basis. As a result there may be some inconsistencies with estimates before and after 2011.

Sex

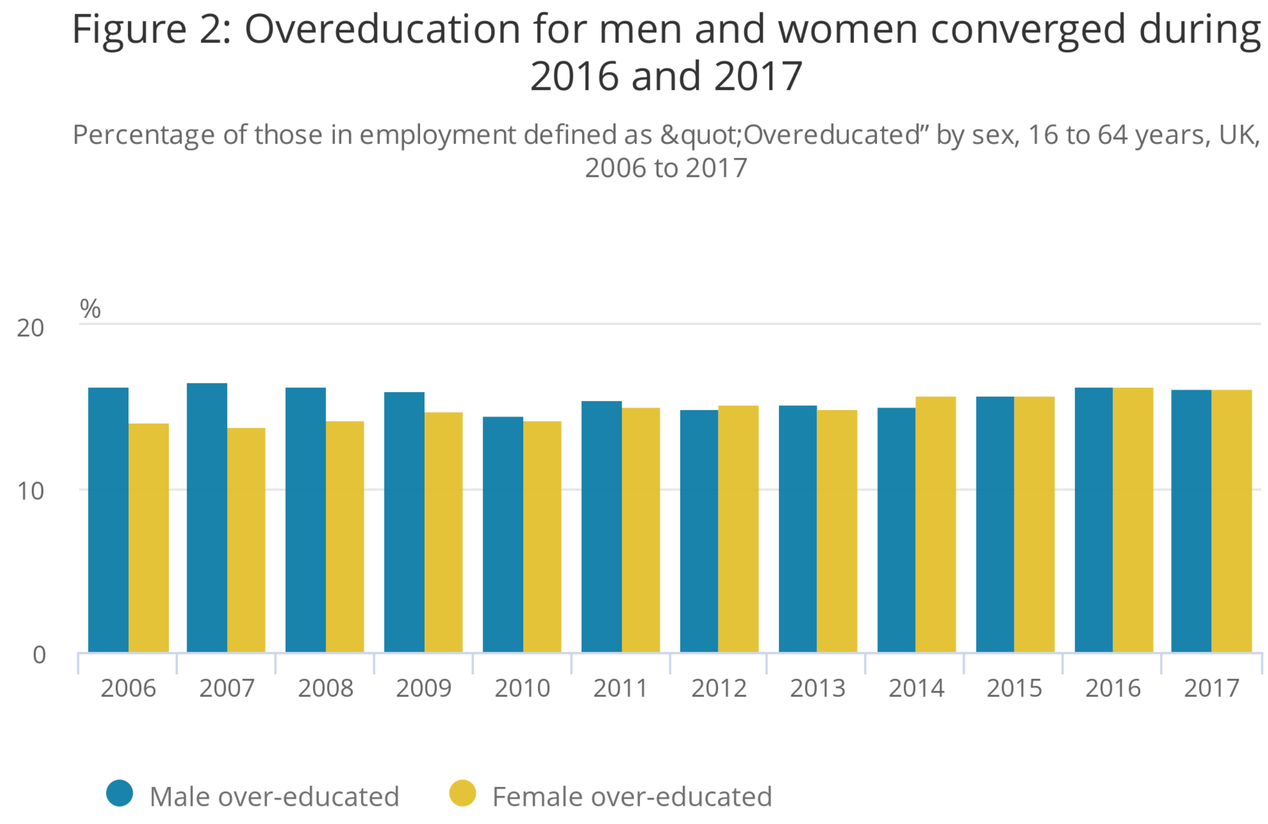
Empirical evidence on the relationship between overeducation and sex has been mixed, with a number of studies concluding that women have a higher overeducation risk than men (Baert and others, 2013; Betti and others, 2011; Ramos and Sanroma, 2013; Tani, 2012; Verhaest and Van der Velden, 2013) as well as those finding no differences between women and men (Blazquez and Budria, 2012; Chevalier, 2003; Chevalier and Lindley, 2009; Frei and Sousa-Poza, 2012). Very few studies indicate that there is higher incidence of overeducation for men (European Commission, 2012; Kiersztyn, 2013).

When we disaggregate our results by sex, we note that from 2006 to 2009 the rate of overeducated males was considerably higher than for females, but this gap started to narrow in 2010 due to an increasing number of women entering higher education (O’Leary and Sloane 2005, Green and Zhu 2010). Our data confirm this trend as from 2006 to 2017 the average education attainment for females increased more compared with males. Females had a slightly lower education attainment in 2006 but this trend was reversed and in 2017 the average education attainment was higher for females.

In 2016 and 2017 the rates for overeducated males and females converged as both sexes became more overeducated compared with previous periods (Figure 2). From 2011 female overeducation fluctuated but generally exhibited a much higher level compared with previous periods, the opposite is true for males.

Figure 2: Overeducation for men and women converged during 2016 and 2017

Percentage of those in employment defined as "Overeducated” by sex, 16 to 64 years, UK, 2006 to 2017



Source: Annual Population Survey – Office for National Statistics

Notes:

1. The data for estimates prior to 2011 were collected on the previous SOC basis (SOC 2000) and have been mapped to an equivalent SOC 2010 basis. As a result there may be some inconsistencies with estimates before and after 2011.

Age

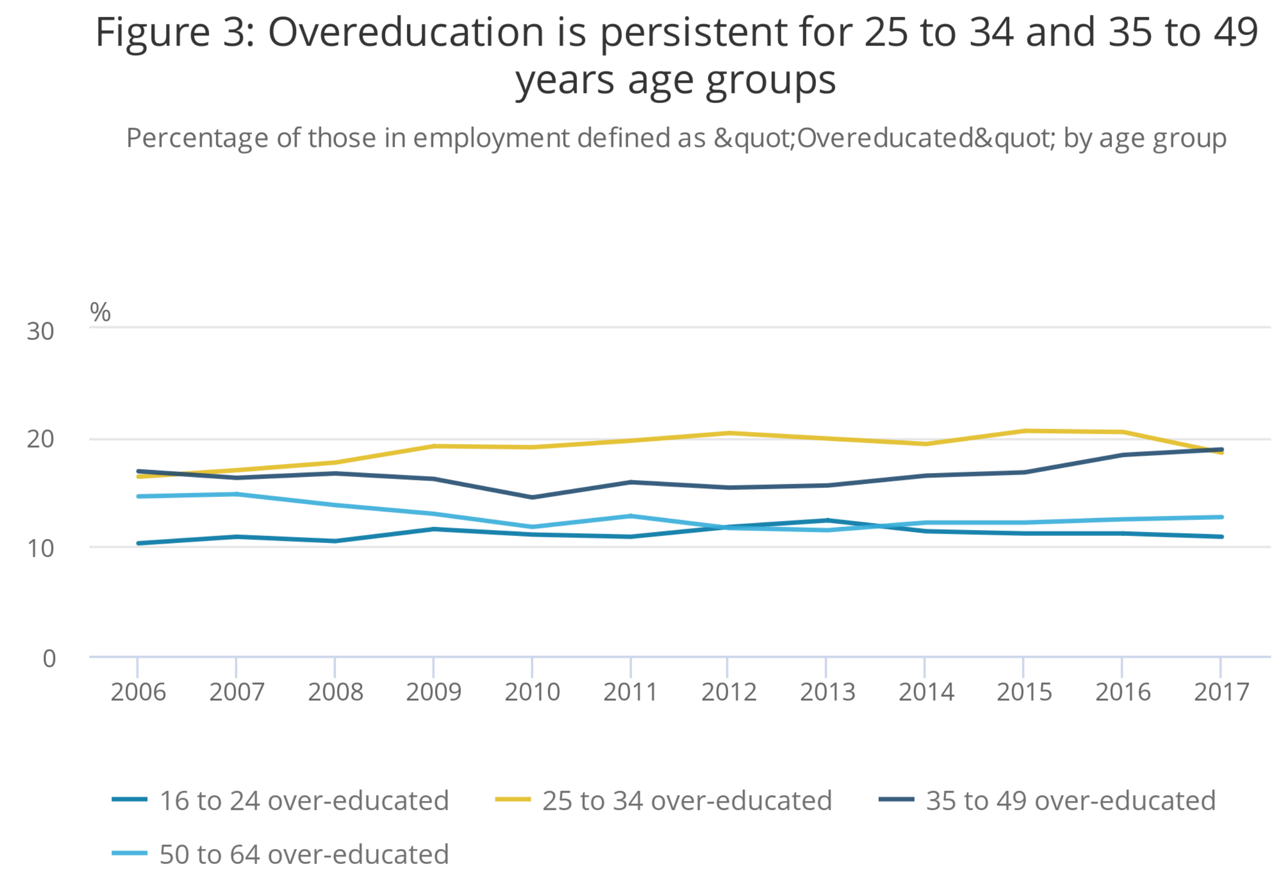
If overeducation was a temporary phenomenon, as predicted by the human capital and the career mobility explanations, we should observe a decline with age. However, the empirical evidence is quite mixed. Some studies emphasise that overeducation decreases with age (Jensen and others, 2010; Robst, 2008; Sutherland, 2012) or that the two have a U-shaped relationship (Tarvid, 2012), while some report that age is irrelevant (Chevalier and Lindley, 2009; Frei and Sousa-Poza, 2012; Kiersztyn, 2013).

Comparing the rate of overeducation for each age group, Figure 3 shows that those aged 25 to 34 years and 35 to 49 years experience the highest rate of overeducation, and the proportion of overeducated workers in the age group 35 to 49 years has increased from 2013 onwards. The finding of higher overeducation in 25 to 34 years age group is consistent with the presence of short-term labour market frictions. However, the high level of overeducation for the 35 to 49 years age group indicates a more persistent phenomenon, particularly as overeducation also features among older workers (50 to 64 years).

Those in employment aged 16 to 24 years experience the lowest rate of overeducation on average. This result should be interpreted with a caveat as the average obtained qualification level for this age group is lower compared with other age groups. For example, almost 70% of those in employment aged 16 to 24 years, hold either GCSE or A level or equivalent as their highest obtained level of education, whereas the corresponding figure for 25 to 34 years age group is 40%. This is an important consideration as education attainment and overeducation are positively correlated. In other words, people who hold GCSE or A level or equivalent education are less likely to be overeducated compared with, for example, graduates.

Figure 3: Overeducation is persistent for 25 to 34 and 35 to 49 years age groups

Percentage of those in employment defined as "Overeducated" by age group



Source: Annual Population Survey – Office for National Statistics

Notes:

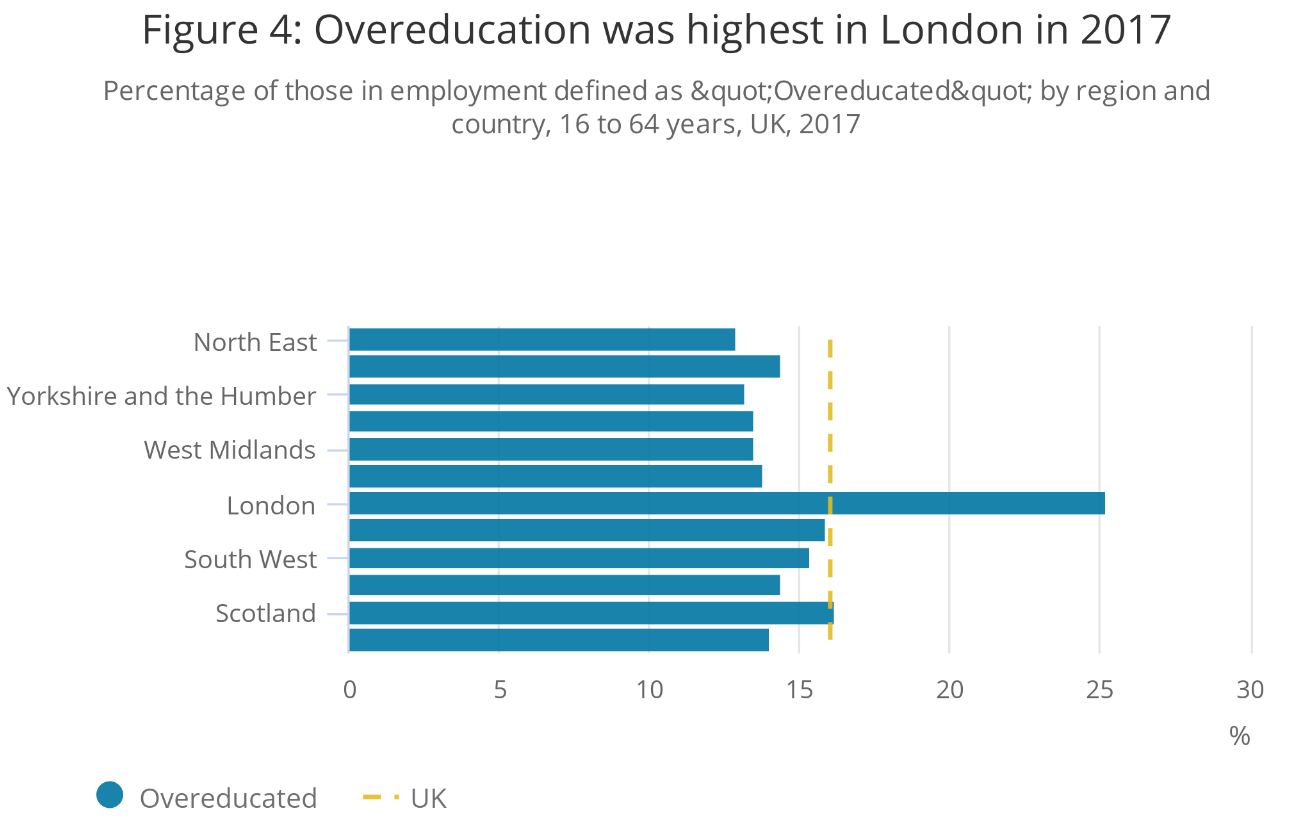
1. The data for estimates prior to 2011 were collected on the previous SOC basis (SOC 2000) and have been mapped to an equivalent SOC 2010 basis. As a result there may be some inconsistencies with estimates before and after 2011.

Region and country

Figure 4 shows that in 2017, London had the highest proportion of overeducated workers in the UK. Results for London are likely to be driven by the composition of the labour force, characterised by a relatively high proportion of immigrants who are typically overeducated. Many foreign nationals working in the UK come to the country to improve their English, hence they may be willing to take a lower-skilled job.

Figure 4: Overeducation was highest in London in 2017

Percentage of those in employment defined as "Overeducated" by region and country, 16 to 64 years, UK, 2017



Source: Annual Population Survey – Office for National Statistics

Notes:

1. Dashed line represents UK average.

Our data show that on average, non-UK born workers tend to be more overeducated compared with the UK-born. The non-UK born overeducated workers accounted for 15.2% of all workers living in London in 2017. It should be noted that the percentage of non-UK born overeducated workers was larger compared with the percentage of overeducated UK-born workers (10%).

Compared with the rest of the country, in 2017, London had the highest proportion of non-UK born matched and non-UK born overeducated and the lowest level of UK-born matched. For the other regions and countries, the average proportion of non-UK born overeducated workers was around 3%.

It should also be noted that our separate analysis shows that half of people living in London between July 2016 and June 2017, aged 21 to 64 years had a level of education above A level standard. This is relatively high compared to other regions and countries and especially compared to the North East where only 30% of people had a level of higher education above A level standard.

First degree subject

This sub-section presents analysis of overeducated graduates whose highest obtained qualification is first degree or equivalent. In 2017, the average overeducation rate for graduates with first degree or equivalent qualification was 30.9%. The overeducation rates for non-recent and recent graduates with first degree or equivalent were 29.2% and 38.6% respectively. This indicates a moderate improvement in the job to education match over time for these groups which is consistent with career mobility proposition.

According to career mobility argument some graduates may begin their careers in a job for which they are overqualified and this job may serve as a stepping stone to a better job in the future. Older graduates who are more likely to possess some pre-study work experience and who have been in their current jobs longer are on average less likely to be overeducated.

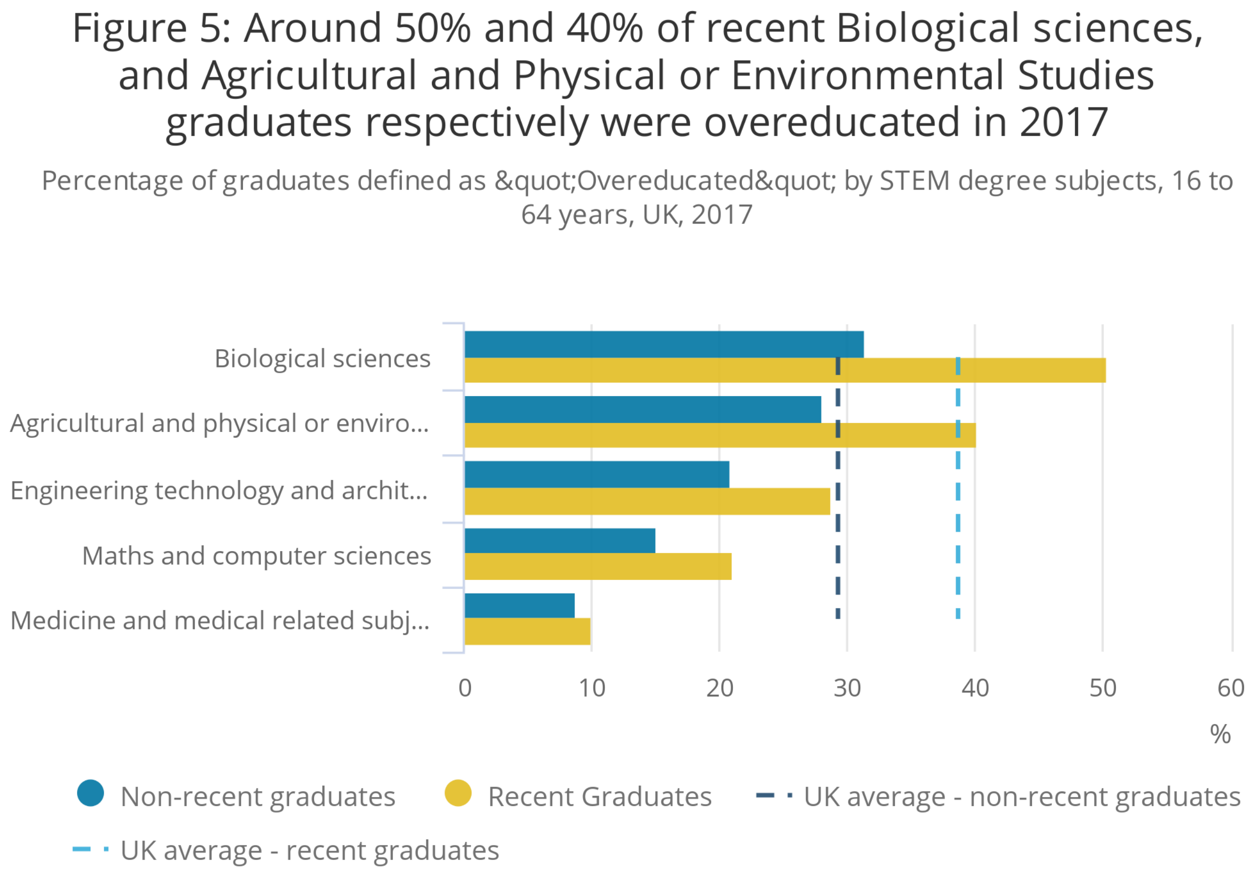
We conduct further analysis of overeducated graduates by graduation cohort. We note that 21.7% of graduates who completed their first degree or equivalent before 1992 were overeducated in 2017, whereas this figure is 23.4% and 24.8% for 1992 to 1999 and 2000 to 2006 cohorts respectively. The corresponding figure is 34.2% for those who graduated during or after 2007. Hence, the incidence of overeducation has increased over time.

Focusing solely on cohort-level overeducation rates is of limited use because of large variation in overeducation rates for graduates across major fields of study. For graduates with first degree or equivalent, the incidence of overeducation varies with respect to the subject studied but graduates from Science, Technology, Engineering and Mathematics (STEM) fields are relatively less likely to end up overeducated within the first five years of completing their degrees.

Figure 5 shows that amongst STEM subjects, the highest incidence of overeducation in 2017 was noted amongst graduates who completed Biological sciences and Agricultural and Physical or Environmental Studies. Around 50% and 40% of recent Biological sciences and Agricultural and Physical or Environmental Studies graduates respectively were overeducated in 2017.

Figure 5: Around 50% and 40% of recent Biological sciences, and Agricultural and Physical or Environmental Studies graduates respectively were overeducated in 2017

Percentage of graduates defined as "Overeducated" by STEM degree subjects, 16 to 64 years, UK, 2017



Source: Annual Population Survey – Office for National Statistics

Notes:

1. Dashed lines represents UK average.

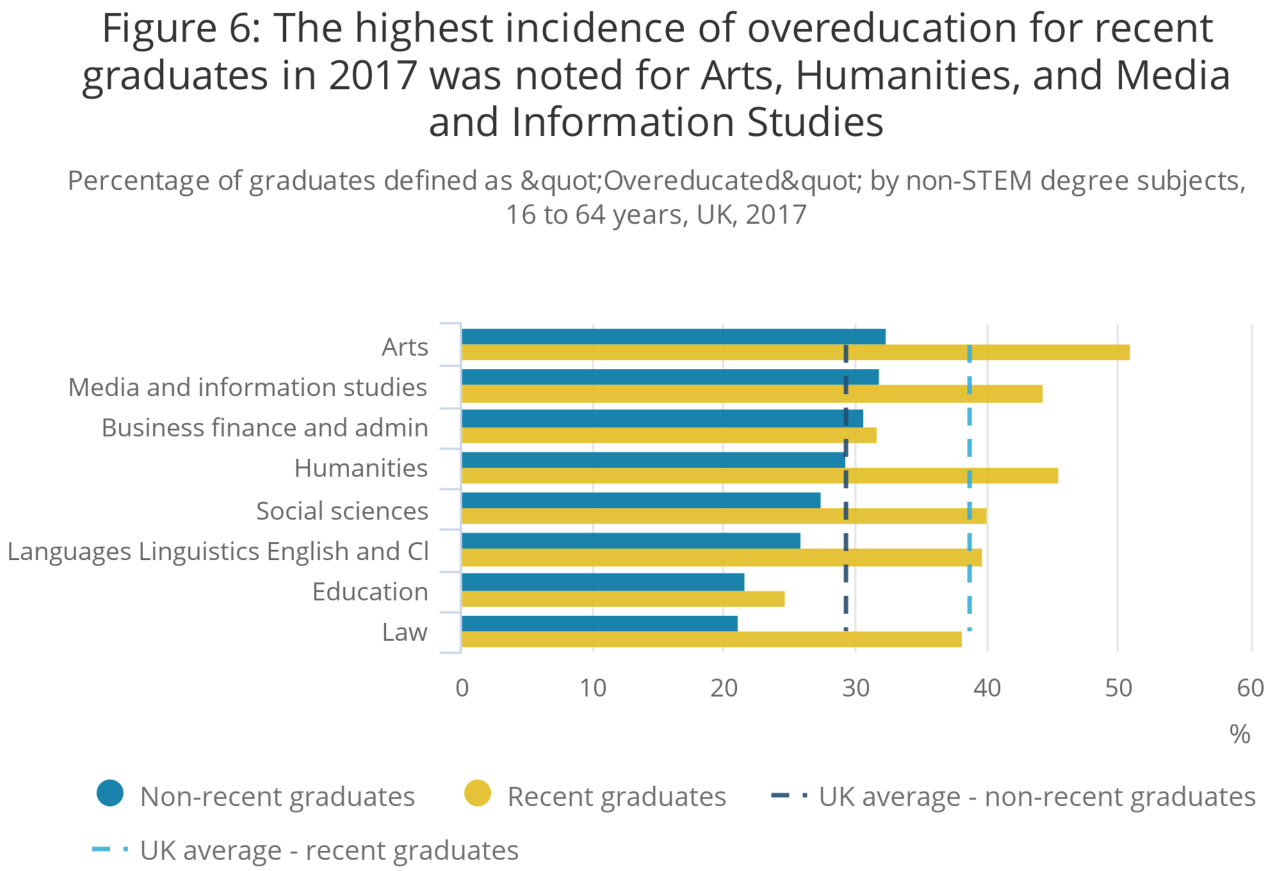
Figure 6 shows that in 2017 amongst recent graduates from non-STEM degree subjects the highest incidence of overeducation was noted for three groups namely: Arts; Humanities and Media and Information studies (51%, 45.5% and 44.4% of overeducated respectively). The four groups of non-recent overeducated graduates belonging to non-STEM degree subject namely Arts; Media and Information Studies; Business, Finance and Administration studies and Humanities had higher than average overeducation rate compared with non-recent graduates overall.

The relatively high rate of overeducation amongst recent Engineering, Technology and Architecture graduates is somewhat surprising. This is possibly due to labour markets for Engineering, Technology and Architecture favouring those with postgraduate qualification, which could result in recently qualified first degree or equivalent holders remaining in lower-skilled work while they attempt to secure employment related to their field of study. The lower rate of overeducation for these graduates after five years suggests that many were able to secure this type of employment or chose to secure skilled employment in other fields.

It should be noted that estimates for medicine and law graduates should be interpreted with caution as law and medicine require a first professional degree, which is normally acquired after a regular first (bachelor's) degree. For the purposes of our analysis which compares graduates with first (bachelor's) degree or equivalent, in this sub-section we exclude those who hold any qualification above that level, that is first professional degree or above in any subject.

Figure 6: The highest incidence of overeducation for recent graduates in 2017 was noted for Arts, Humanities, and Media and Information Studies

Percentage of graduates defined as "Overeducated" by non-STEM degree subjects, 16 to 64 years, UK, 2017



Source: Annual Population Survey – Office for National Statistics

Notes:

1. Dashed lines represents UK average.

Next, we aim to assess whether overeducated workers may be at a disadvantage compared to their matched counterparts. In particular, we aim to assess whether overeducation may be associated with lower earnings and whether any observed effect may differ for men and women and for recent graduates.

Notes for: Results-descriptive analysis

1. A graduate is defined as a person who is in employment, aged between 16 and 64 years, not enrolled on any educational course and who has a level of higher education of first degree or equivalent level standard. Following qualifications are defined as first degree or equivalent: (1) NVQ level 5 (2) Level 8 Diploma (3) Level 8 Certificate (4) Level 7 Diploma (5) Level 7 Certificate (6) Level 8 Award (7) First degree/foundation degree (8) Other degree.

4 . The relationship between wages and overeducation

To investigate the relationship between wages and overeducation, we estimate the earnings difference between overeducated and matched individuals while controlling for relevant individual and job characteristics. There is a large empirical literature on the impact of the education mismatch on productivity, generally captured through average hourly wages. This body of evidence draws upon the human capital theory, and on the assumption that wages equal marginal productivity in a competitive market.

Our indicators of required, over and deficit education are based on the distribution of actual educational attainments in each occupation defined at the three-digit SOC level and constructed by two age groups 16 to 35 years and 36 to 64 years. Our measure classifies as overeducated a worker whose education (years of schooling) deviates positively from the observed occupation average by more than one standard deviation, and as undereducated those with a negative deviation from the mean in excess of one standard deviation.

We also construct a variant of this procedure, replacing mean-centred bracket with a different measure of central location, the distribution mode. We use both measures (mean and mode) as robustness checks in our regression analysis in models 1 to 4 (Table 1). We control for various individual and job characteristics which may impact both the incidence of overeducation as well as wages. Full information can be found in Appendix 1.

Most of the empirical literature on education mismatch support the following two hypotheses:

- H1: Overeducated workers earn more than their co-workers who are matched but less than workers with the similar level of education in matched jobs; and
- H2: Under-educated workers earn less than their co-workers who are matched but more than similarly educated workers employed in lower level jobs.

Two main specifications have been used, where the first one represents extended versions of the Mincer (1974) equation:

$$\ln(w) \ = \ x\beta_1 \ + \ \beta_2 S^R \ + \ \beta_3 S^O \ + \ \beta_4 S^U \ + \ u_i$$

Where ln(w) is a natural log of hourly wages, x is a vector of workers’ and job characteristics, S^R is the number of years of required schooling, S^O is the number of years of over-education and S^U is the number of years of deficit schooling (undereducation).

An alternative specification uses dummy variables to identify workers with different education level following Verdugo and Verdugo (1998):

$$\ln(w) \ = \ x\beta_1 \ + \ \beta_2 E \ + \ \beta_3 D^O \ + \ \beta_4 D^U \ + \ u_i$$

Where ln(w) is a natural log of hourly wages, E is the number of years of education for each worker, D^O and D^U are the dummy variables capturing whether the worker is over or under educated for his or her occupation. This specification compares overeducated and undereducated workers with those who are matched and have similar level of education.

We also conduct quantile regression analysis for graduates where instead of considering the effect of overeducation on the mean wages, we look at 25th percentile, median and 75th percentile of wages. We aim to assess whether the effect of overeducation differs across different points of the pay distribution.

The analysis will be carried out for the whole sample and for a sub-sample of graduate workers. The latter will include controls for the degree subject and will distinguish between recent graduates (first degree completed within the past five years) and non-recent graduates.

Table 1: OLS regression, hourly wage estimation results, 2017

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5 (Female)**	**Model 6 (Male)**
Variables	**log_hourpay**	**log_hourpay**	**log_hourpay**	**log_hourpay**	**log_hourpay**	**log_hourpay**
Obtained education	0.084***	0.076***				
	(0.003)	(0.003)				
Obtained education squared	-0.002***	-0.001***				
	(0.000)	(0.000)				
Overeducated dummy (mean)	-0.081***					
	(0.008)					
Undereducated dummy (mean)	0.058***					
	(0.008)					
Overeducated dummy (mode)		-0.033***				
		(0.007)				
Undereducated dummy (mode)		0.024***				
		(0.009)				
Required education years (mean)			0.100***		0.082***	0.111***
			(0.003)		(0.004)	(0.004)
Overeducation years (mean)			0.013***		0.010***	0.017***
			(0.002)		(0.002)	(0.002)
Undereducation years (mean)			-0.100***		-0.106***	-0.089***
			(0.006)		(0.008)	(0.009)
Required education years (mode)				0.031***		
				(0.001)		
Overeducation years (mode)				0.026***		
				(0.001)		
Undereducation years (mode)				-0.082***		
				(0.005)		
Constant	0.289***	0.396***	-0.299***	0.724***	0.058	-0.552***
	(0.051)	(0.051)	(0.057)	(0.047)	(0.082)	(0.081)
Observations	76,360	76,360	76,337	76,337	40,342	35,995
R-squared	0.432	0.430	0.433	0.424	0.412	0.426

Notes

1. Standard errors in parentheses. Back to table

2. *** p<0.01, ** p<0.05, * p<0.1. Back to table

Table 1 shows coefficient estimates based on the two specifications of the wage equations. As a robustness check, we use measures of educational mismatch based on the mean and on the mode years of required education in each occupation. The main findings can be summarised as the following.

Models 1 and 2 compare overeducated and undereducated employees with those who are matched to their jobs and have similar level of education. Both models predict a penalty for overeducation and a premium for undereducation. The penalty is larger when using a definition of educational mismatch based on the average (mean) years of education (8.1%) as opposed to a mode-based measure (3.3%). This means that the wage of an overeducated worker is between 3.3% and 8.1% lower compared to the wage of a worker with a similar level of education who is matched to the right occupation.

Models 3 and 4 compare overeducated and undereducated employees with their matched co-workers who may not have the same level of education. To correctly interpret the results, we express the total number of years of education (E) as a linear combination of required (E_r), over (E_o) and undereducation (E_u): $E = E_r + E_o – E_u$. The coefficient for required years of education is 0.100-implying that an additional year of required education increases wages by 10% – while the coefficient for surplus education is 0.013. Hence, although there is a positive return on wages for overeducated employees, the size of the coefficient indicates that for years of education beyond required a worker receives 8.7% wage penalty (0.100-0.013=0.087). When measuring education mismatch using the mode, returns on wages from an additional year of required education are lower than expected (0.031) and the corresponding coefficient for overeducation is 0.026. This means that education beyond what is required receives only 0.5% wage penalty. These results indicate that mode-based measures induce lower penalty for overeducation compared with mean-based measures.

Models 5 and 6 investigate the impact of overeducation on wages by estimating two separate models for female and male employees respectively. Given that the literature provides stronger support for the extended Mincer equation approach (Hartog 2000), we carry on with this specification in the rest of our analysis. Although the prevalence of overeducation is similar for women and men in 2017 (Figure 2), our estimates suggest that the diminishing return on wages is somewhat more pronounced for overeducated men than for women. Women receive 7.2% (0.082-0.010=0.072) lower return on wages for education beyond required level, whereas men receive 9.4% (0.111-0.017=0.094) lower return on wages for education beyond required level. However, women seem to receive higher penalty for undereducation. We should also note that the return on wages from the required education for men is higher compared to women as men earn 11.1% return on additional year of required education, whereas women earn 8.2%. This means that after controlling for individual and job characteristics, our results show that men receive higher returns on wages for required education.

Apart from education, some other individual characteristics feature prominent differences between the two sexes. The results for both individual and job controls are presented in Table 6, Appendix 2.

Evidently, the household context has different implications for men and women. Mothers earn less than childless females, whereas fathers do not suffer any fatherhood penalty. In fact, fatherhood has a significant and positive effect on wages. This is important to note because females' decisions and subsequently earnings may be to a greater extent (compared with males') driven by personal traits which we cannot observe in our data.

The estimators indicate that males' wages benefit from having dependent children. The results for men show that married men earn more than unmarried men and fathers earn more than childless men. By contrast, having dependent children lowers females' income. In other words, household characteristics that stimulate breadwinner role seem to induce wage premium only for men whereas women do not benefit.

We should also note that overeducation does not seem to make significant contribution to the gender wage gap. This is because, as shown in descriptive analysis, females are overeducated almost as often as males and given that overeducation is more severely penalised for men than for women.

Our findings indicate that a noticeable part of the wage gap may be attributed to women's labour market decisions as only males are granted breadwinner wage premiums, whereas women suffer wage reductions when they have dependent children.

In conclusion, although the overall returns on wages from investment in education are positive, overeducated employees earn lower wages than those with similar education who are matched to their jobs. In other words, there are positive but diminishing returns to overeducation. Undereducated workers earn more than similarly educated workers who are employed in lower-level jobs, but less than their co-workers whose level of education corresponds to that required in their occupation. These findings support hypotheses 1 and 2.

Graduates

In this sub-section, we aim to evaluate whether the impact of overeducation differs depending on the type of degree and whether younger (recent) graduates earn higher or lower wages compared to non-recent graduates (Table 2).

Table 2: OLS regression, hourly wage estimation results for graduates, 2017

	Model 7	**Model 8**	**Model 9**	**Model 10**	**Model 11**
Variables	**log_hourpay**	**log_hourpay**	**log_hourpay**	**log_hourpay**	**log_hourpay**
Required education	0.054***	0.054***	0.054***	0.053***	0.054***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Overeducation	-0.014***	-0.018***	-0.022***	-0.018***	-0.017***
	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
Recent graduate	-0.005	-0.003	-0.005	-0.016	-0.005
	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)
STEM degree subject	0.021**	0.021**	0.021**	0.021**	0.016*
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Age	0.064***	0.065***	0.064***	0.063***	0.064***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female	0.018	0.017	0.014	0.019	0.017
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Tenure	0.010***	0.010***	0.010***	0.010***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Overeducation*Tenure		0.000*			
		(0.000)			
Overeducation*Female			0.014***		
			(0.005)		
Overeducation*Recent graduate				0.012**	
				(0.005)	
Overeducation*STEM degree subject					0.011*
					(0.005)
Constant	-0.024	-0.027	-0.010	0.008	-0.024
	(0.141)	(0.141)	(0.141)	(0.142)	(0.141)
Observations	17,672	17,672	17,672	17,672	17,672
R-squared	0.403	0.403	0.403	0.403	0.403

Source: Annual Population Survey – Office for National Statistics

Notes

1. Robust standard errors in parentheses. Back to table
2. *** p<0.01, ** p<0.05, * p<0.1. Back to table

In all specifications, we find that overeducated graduates experience negative returns to overeducation, with an effect ranging between negative 0.014 and negative 0.022. There are no statistically significant differences in wages between recent and non-recent graduates, nor by gender. However, the type of degree does make a difference and all models predicts higher wages for workers with STEM degrees, compared to non-STEM. Other key findings include the following.

The wage penalty for graduates decreases the longer they stay with their employer, as both the coefficient on tenure and the interaction between tenure and overeducation are positive and statistically significant, although the latter effect is very small. This suggests that graduates may substitute education for the lack of previous work experience, accepting jobs requiring less education than they possess to acquire the necessary experience. It may also indicate that overeducation may not have negative effect on graduates' productivity as they gain more experience. Overall, increased tenure is rewarded by employers.

We investigate whether the negative returns on wages for overeducated graduates differ by sex. Our prior expectation was that negative returns on wages may be driven by the increasing number of overeducated female graduates. In contrast, we find that overeducated female graduates bear lower penalty (negative 0.008) compared with men (negative 0.022).

Another important question is whether recent graduates, that is those who graduated within the past five years, pay a higher penalty for overeducation compared to non-recent graduates. The intuition is partly driven by studies showing that managers are more inclined to hire more experienced as opposed to younger workers because the former are perceived as being more reliable and professional (Corgnet and others 2015). As a consequence, younger graduates are expected to suffer higher penalty for overeducation. However, we find that the negative effect of overeducation for non-recent graduates (negative 0.018) is nearly offset by the impact of overeducation for younger (recent) graduates (0.012). In other words, although recent overeducated graduates still have a negative return on wages, they earn around 1.2% more compared with non-recent overeducated graduates. Contrary to existing findings (Frenette (2004), Mavromaras and others (2010) and Carroll and Tani (2013)), recent graduates in the UK are penalized for their overeducation, although the penalty is lover compared to non-recent graduates. This suggests that recent graduates have specific skills or unobservable characteristics that are better valued in the labour market compared with non-recent graduates.

We look at the impact of overeducation by type of degree by including the interaction between a dummy indicating a degree in STEM subjects and overeducation. Results show that a degree in STEM subjects contributes towards reducing the penalty for overeducation from negative 0.017 (for non-STEM degrees) to negative 0.006.

An important information that is missing from our data is a direct measure of individuals' cognitive and non-cognitive skills (personality traits), both considered to be important for a successful labour market experience (Heckman and others 2006). Although education is generally used as a proxy for cognitive abilities, the fast expansion of the higher education in UK since the 1980s might have compromised quality for quantity (Chevalier and Lindley 2009).

This means that graduates may lack the adequate skills for the job despite having more education than required for their occupation (that is they might not be over-skilled). Empirical research using different proxies show that graduates with lower ability face a higher risk of overeducation (Barone and Ortiz 2011; Chevalier 2003; Lianos et al. 2004; Tarvid 2012; Verhaest and Omey 2010) and personality traits may be more important than ability in determining overeducation (Blazquez and Budria 2012; Tarvid 2013).

The nature of our data does not allow us to control for individuals' ability traits in regression analysis. However, if our results were affected by graduates' abilities we would expect the wage penalty on overeducation to be higher for low earners but much smaller or non-existent for higher earners. To test this proposition, we conduct a quantile regression analysis, which allows us to check how the impact of overeducation (and other controls) varies along the income distribution.

Table 3 shows that there are increasing returns to required education as we move up towards better paid jobs. An additional year of required education increases wages by 4% in the bottom quantile and by 6.1% in the top quantile. However, the penalty for overeducation changes marginally, from 1% to 1.1%. There is no wage difference between recent and non-recent graduates, while graduates with STEM degrees earn significantly higher wages only in the bottom half of the wage distribution. In higher-level occupations the type of degree does not significantly affect wages. This indicates that individuals in highly-paid jobs are endowed with additional unobserved skills, which are independent of the degree subject.

Table 3: Quantile regression, hourly wage estimation results for graduates, 2017

	Model 12	**Model 13**	**Model 14**
Variables	**q25**	**q50**	**q75**
Required education	0.040***	0.045***	0.061***
	(0.005)	(0.004)	(0.004)
Overeducation	-0.010***	-0.011***	-0.011***
	(0.002)	(0.003)	(0.003)
Recent graduate	-0.009	-0.013	-0.004
	(0.013)	(0.012)	(0.016)
STEM degree subject	0.016*	0.011**	0.007
	(0.008)	(0.005)	(0.006)
Age	0.047***	0.052***	0.065***
	(0.002)	(0.003)	(0.004)
Constant	0.291**	0.392***	0.156
	(0.124)	(0.123)	(0.117)
Observations	17,672	17,672	17,672

Source: Annual Population Survey – Office for National Statistics

Notes

1. Standard errors in parentheses. Back to table
2. *** p<0.01, ** p<0.05, * p<0.1. Back to table

Notes for: The relationship between wages and overeducation

1. Education mismatch refers to the difference between the worker's attained level of education and the education required on the job (Hartog 2000). If we define as E the number of years an employee has invested in education, Er the years of education required in an occupation, the mismatch arises when either E>Er, (overeducation), or E<Er (undereducation) (Cohn and Khan 1995).

2. The Mincer earnings function is a single-equation model that explains wage income as a function of schooling and experience, named after Jacob Mincer.

3. The dependent variable (hourly wages) is expressed in log form. If the distribution of a variable has a positive skew, taking a natural logarithm of the variable helps fitting the variable into a model. Also, when a change in the dependent variable is related with percentage change in an independent variable, or the other way around, the relationship is better modelled by taking the natural log of either or both variables.

4. We construct both mean and mode measures of education mismatch (over and under education) which we use in regression analysis (Models 1 to 4, Table 1). We describe our approach in more detail in Appendix 1.

5 . Conclusion

This article investigates the incidence of overeducation by sex, age and region and the incidence and persistence of overeducation for graduates by type of first degree. We also investigate the relationship between overeducation and wages for males and females and for recent and non-recent graduates.

We note that the incidence of overeducation does seem higher for certain age groups and in particular for those aged 25 to 34 years and 35 to 49 years. The relatively high incidence of overeducation for 35 to 49 years age group indicates that overeducation is a persistent phenomenon in the UK labour market. The incidence of overeducation is also higher in London compared with other UK regions and countries. It is higher for certain first degree subjects but it is generally lower for graduates with Science, Technology, Engineering and Mathematics (STEM) degrees, whereas the results for females and males in the most recent period (2017) do not seem to differ.

When we compare the relationship between overeducation and wages for men and women, we find that the wage penalty for overeducation is somewhat higher for men compared with women. Given that the incidence of overeducation does not differ between men and women and that men on average bear higher penalty for overeducation, we conclude that overeducation does not seem to contribute to the gender wage gap. Instead the gender wage gap may be attributed to women's labour market decisions and traditional gender roles as only males are granted breadwinner wage premiums, whereas women suffer wage reductions when they have dependent children.

In relation to graduates, we argue that overeducation would not be a concern if it was strictly a short-term phenomenon. However, our results show that a non-negligible number of graduates are still overeducated five years after completing their first degree (29.2%). With regards to the effect of overeducation on wages, recent (younger) overeducated graduates had a lower penalty on overeducation compared with older (non-recent) overeducated graduates. This finding indicates that younger graduates may have characteristics which are valued in the labour market. In addition, our analysis shows that earnings of overeducated STEM subject graduates contribute towards reducing the wage penalty for overeducation and almost offsetting it.

However, we do acknowledge that older (non-recent) graduates may have different characteristics compared with younger (recent) graduates in terms of abilities and/or motivation. In this respect our data has limitations as it does not allow us to capture these effects. Hence, we conduct quantile regression analysis under the assumption that if the wage penalty amongst graduates was related to lower ability, the penalty would be more prominent among the graduates who are lower earners. We find no support for lower ability argument as we observe no significant differences across the wage distribution for overeducated graduates.

However, we cannot rule out the possibility that for some graduates, overeducation and associated lower earnings may simply be a matter of choice and personality traits. In other words, some graduates may willingly refrain from maximizing their individual income due to hidden preferences (Frank, 1978). It follows that to fully depict the effect of overeducation on workers' productivity it is not sufficient to concentrate exclusively on earnings. Hence, we propose to extend our analysis to investigate the effect of overeducation on productivity directly, measured in terms of output per worker or total factor productivity.

Limitations

Our analysis investigates the relationship between overeducation and hourly wages. We acknowledge, however, that overeducation may be associated with some other individual or societal cost and/or benefits in addition to its effect on wages (Green and Henseke, 2016). These may relate to positive externalities such as lower crime rates and knowledge spill-overs or costs to the taxpayer, but the assessment of these is not within the scope of our present analysis.

Our data have several important limitations. To compute estimates of overeducation and conduct hourly wages analysis we use the Annual Population Survey (APS) (approximately 70,000 observations per year). Compared with the Labour Force Survey (LFS), the APS provides a larger sample. The APS is a sample survey and all estimates from it are subject to sampling variability. Sampling variability is dependent on several factors, including the size of the sample, the size of the estimate as a percentage of the population and the effect of the design of the sample on the variable of interest. Therefore, it is subject to a margin of uncertainty, as different samples provide different results.

In the LFS, earnings information is asked in the first and fifth wave and an hourly earnings variable is created using information on both pay and hours. Since the APS includes wave one and five from the LFS as well as regional boosts, the APS earnings question will be asked of all eligible employees. This is not calculated for those in self-employment. Gross hourly earnings data are known to be underestimated in the LFS or APS and this is principally because of proxy responses. To correct for this, our regression analysis includes a dummy control variable for proxy responses.

Hence, the Annual Survey of Hours and Earnings (ASHE) is the preferred source for earnings information. The ASHE is collected from the employer and as such the earnings information is thought to be more reliable as it is mainly provided with reference to company records. In contrast, the APS wages data is provided by the individual and it is subject to recall error, which is compounded when information is provided by proxy response.

It should be noted however, that the LFS and APS, unlike ASHE, allow us to investigate the relationship between overeducation and wages since both surveys contain sufficient information on wages as well as other individuals' characteristics such as the level of education.

Another important limitation of the data is that they do not have a sufficient panel element. The use of cross-section data in regression analysis makes it difficult to infer causality. This is because both education and overeducation are potentially endogenous. However, comparisons between males and females and between recent and non-recent graduates are made on the assumption that the sources of endogeneity and bias are similar across these categories and that therefore the differences in the estimates are informative.

6 . Authors

Maja Savic, Michela Vecchi and Adama Lewis

7 . Acknowledgements

We would like to thank Professor Rob Wilson at the Institute for Employment Research, University of Warwick for providing us with SOC to O*NET mapping document.

8 . References

Baert, S, Cockx, B, Verhaest, D (2013), Overeducation at the start of the career: stepping stone or trap? Lab. Econ. 25, pages 123 to 140.

Barone, C, Ortiz, L (2011). Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation. ‘Higher Education’, 61(3), pages 325 to 337.

Betti, G, D’Agostino,A, Neri, L. (2011), Educational mismatch of graduates: a multidimensional and fuzzy indicator. Soc. Indic. Res. 103(3), pages 465 to 480.

Blázquez, M, Budría, S. (2012), Overeducation dynamics and personality. Educ. Econ. 20(3), pages 260 to 283.
Carroll, D, Tani, M (2013). Over-education of recent higher education graduates: New Australian panel evidence. ‘Economics of Education Review’, 32, pages 207 to 218.

Chevalier, A (2003), Measuring over-education. Economica 70(279), pages 509 to 531 (2003).

Chevalier, A, Lindley, J (2009), Overeducation and the skills of UK graduates. J R Stat Soc: Series A (Statistics Society) 172(2), pages 307 to 337.

Cohn, E, Khan, S. P. (1995), The wage effects of overschooling revisited. ‘Labour Economics’, 2(1), pages 67 to 76.

Dang, A, Nguyen, N, Nguyen, L. (2017), The Wage Impacts of Over-education in Nine Major Occupation Groups in the UK: Propensity Matching Score Approach.

European Commission (2012), DG Employment, Social Affairs and Inclusion: Employment and social developments in Europe 2012. Tech. rep.

Forrest, R. (2015), Housing and family wealth, International Journal of Housing Policy, 15:4, pages 491 to 494.

Frei, C., Sousa-Poza, A.(2012), Overqualification: permanent or transitory? Appl. Econ. 44(14), pages 1837 to 1847.

Frenette, M. (2004). The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification. ‘Economics of Education Review, 23’(1), pages 29 to 45.

Green, F (2016). Skills demand, training and skills mismatch: a review of key concepts, theory and evidence. ‘UK Foresight Future of Skills and Lifelong Learning Project’.

Green, F, Henseke, G. (2016), Should governments of OECD countries worry about graduate underemployment?. ‘Oxford Review of Economic Policy’, 32(4), pages 514 to 537.

Green, F, Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. ‘Oxford economic papers’, 62(4), pages 740 to 763.

Groot, W, and Van Den Brink, H. M. (2000). Overeducation in the labor market: a meta-analysis. ‘Economics of education review, 19’(2), pages 149 to 158.

Hartog, J (2000), Over-education and earnings: where are we, where should we go?. ‘Economics of education review, 19’(2), pages 131 to 147.

Heckman, J. J, Stixrud, J, Urzua, S (2006), The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. ‘Journal of Labor economics, 24’(3), pages 411 to 482.

ILO, (2014), Skills Mismatch in Europe.

Jensen, U, Gartner, H, Rässler, S(2010), Estimating German overqualification with stochastic earnings frontiers. AStA Adv. Stat. Anal. 94(1), pages 33 to 51.

Kiersztyn, A (2013), Stuck in a mismatch? The persistence of overeducation during twenty years of the post-communist transition in Poland. Econ. Educ. Rev. 32, pages 78 to 91.

Lianos, T. P, Asteriou, D, Agiomirgianakis, G. M (2004). Foreign university graduates in the Greek labour market: Employment, salaries and overeducation. ‘International Journal of Finance & Economics, 9’(2), pages 151 to 164.

Mavromaras, K, McGuinness, S, O'leary, N, Sloane, P, Fok, Y. K. (2010). The problem of overskilling in Australia and Britain. ‘The Manchester School, 78’(3), pages 219- to 241.

McGuinness, S (2006). Overeducation in the labour market. ‘Journal of economic surveys, 20’(3), pages 387 to 418.

McGowan, M A., and Andrews, D (2015). Labour market mismatch and labour productivity.

Mincer, J, 1974. Schooling, Experience, and Earnings. Columbia University Press, New York.

O’Leary, N. C, Sloane, P. J. (2005). The return to a university education in Great Britain. ‘National Institute Economic Review’, 193(1), pages 75 to 89.

Ramos, R, Sanromá, E (2013), Overeducation and local labour markets in Spain. Tijdschr. Econ. Soc. Ge. 104 (3), pages 278 to 291.

Robst, J (2008), Overeducation and college major: expanding the definition of mismatch between schooling and jobs. Manch. Sch. 76(4), pages 349 to 368.

Sutherland, J (2012), Qualifications mismatch and skills mismatch. Educ. Train. 54(7), pages 619 to 632. Tani, M (2012), Does immigration policy affect the education-occupation mismatch? Evidence from Australia. Australian Bull. Lab. 38(2), pages 111 to 140.

Tarvid, A (2012), Effects of the field of higher education on the prevalence of over-education in European countries. In: Kesner, J. (ed.) Education: Evaluation, Reform and Policy, pages 167 to 184. Athens Institute for Education and Research, Athens.

Tarvid, A (2013), Unobserved heterogeneity in overeducation models: Is personality more important than ability?. Procedia Economics and Finance, 5, pages 722 to 731.

Verdugo, R. R, and Verdugo, N. T (1989). The impact of surplus schooling on earnings: Some additional findings. 'Journal of Human Resources', pages 629 to 643.

Verhaest, D, Omey, E (2010), The determinants of overeducation: different measures, different outcomes?. 'International Journal of Manpower, 31'(6), pages 608 to 625.

Verhaest, D, Van der Velden R (2013), Cross-country differences in graduate overeducation. Eur. Sociol. Rev. 29 (3), pages 642 to 653.

9 . Appendix 1

Method

Measures of micro-level education mismatch

The concept of overeducation (undereducation) means having more (less) education than required for the job, but uniquely accepted typology or measurement framework is lacking. Three different approaches exist (Table 4) with their own advantages and disadvantages. Moreover, different measures may lead to different results in terms of both estimating the prevalence of education mismatch as well as its relationship with wages.

The measurement of education mismatch has made use of both subjective and objective measures. All measures have some shortcomings and may be prone to a biased evaluation of the extent of over and under education. Common to all measures, the analysis starts with the evaluation of the level or years of education required in each occupation, usually following the three-digit Standard Industrial Occupation (SOC) codes. Levels or years of education above the required are defined as overeducation, while levels or years below are defined as undereducation. Objective measures can be divided into two sub-groups, namely normative and statistical.

Table 4: Measures of education mismatch

Measure	Description	Advantages	Disadvantages
Normative (objective)	Use a pre-determined mapping between the job and the required education level	Easily measurable	Assumes constant mappings over all jobs of a given occupation
		Objective	Costly to create and update mapping
Statistical (objective)	The are those with education level higher by some ad-hoc value than the mean or mode of the sample within a given occupation	Easily measurable	Assumes constant mappings over all jobs of a given occupation
		Objective	Sensitive to cohort effects
		Always up-to-date	Results depend on the level of aggregation of occupations
Self-assessment (subjective)	Respondents are asked about their perceptions of the extent their education or skills are used in their job	Always up-to-date Corresponds with requirements in the individual firm	Subjective bias: respondents may overstate job requirements, inflate their status or reproduce actual hiring standards

Source: Authors' elaboration: (Hartog 2000)

Normative measure is based on experts' views, who specify the required level of education for the job titles in an occupational classification. One example is O*NET (Occupational Information Network), a unique, comprehensive database of worker competencies, job requirements, and resources. O*NET replaces the Dictionary of Occupational Titles (DOT) and it is the primary source of occupational information in the US.

The O*NET database contains a wide variety of worker and job oriented data categories as well as a rich set of variables that describe job and worker characteristics, including skill and educational requirements. The O*NET-SOC Occupation Taxonomy covers work performed in the US economy and defines the set of occupations for which data is collected. There is no dataset of this kind currently available in the UK.

Statistical measures are based on realized matches. Required education is computed using the mode or the mean level or years of education within each SOC classification. When using the mean, workers are classified as over or undereducated if their level of education is more or less than one standard deviation above their occupations' mean education level.

The statistical method assumes that the mean level of educational attainment represents the requirement for the occupation. It should be acknowledged, however, that educational attainment does not fully capture the skills required for each type of job for example experience, on-the-job training, non-exam based learning and some vocational qualifications. However, the approach does benefit from being measurable and generally comparable over time.

Subjective measures are derived from workers self-assessment via employees' surveys. These measures may be prone to measurement error due to respondents' subjectivity bias as respondents' may overstate job requirements or inflate their status within the company.

While previous meta studies of education mismatch literature by Groot and van den Brink (2000) and Rubb (2004) indicate the extent to which various definitions tend to identify different people as being either overeducated, undereducated or matched and generate different estimates of the incidence of and returns to overeducation; they provide no indication of which measure is closest to the true incidence or the extent to which any particular approach generates biased estimates. To a large extent, the level of correlation is likely to vary according to the dataset being used by researchers and the institutional or economic arrangements of the country in question.

The issue of empirical bias associated with the various definitions was addressed by Groot and van den Brink (2000) who conducted a cross-country meta-analysis of 25 studies, utilizing the various subjective and objective methodologies. The authors found that the standard deviation-based measure (statistical measure) tended to yield the lowest estimate of the incidence of overeducation.

The finding of a lower incidence under the standard deviation approach is not surprising, as the methodology requires education levels to be at least one standard deviation above the mean before overeducation is determined whilst the other approaches have no such requirement. In relation to the wage equation meta-analysis, the authors did not find any of the methodological approaches to significantly influence estimated returns.

A meta-analysis by Rubb (2004) found that neither the subjective nor occupational dictionary approaches (normative measure) yielded estimates of the overeducation wage effect that were significantly different from a measure based on the mean occupational level (statistical measure). Thus, one might conclude from such cross-country studies that whilst there are serious concerns relating to the low correlation between the various measures of overeducation the evidence would suggest that, in terms of estimating the returns to overeducation, the various approaches generate broadly consistent evidence. In section four we test this proposition and estimate the relationship between over and under education and wages using both mean and mode based measures of education mismatch.

Statistical method

We have previously published Analysis of the UK labour market-estimates of skills mismatch using measures of over and under education: 2015. In our previously published article as well as in our present analysis, we broadly follow ILO (2013, 2014) statistical approach. In our descriptive analysis we use highest qualification or trade apprenticeship as a proxy for educational attainment and job requirement, instead of years of full-time education.

When we estimate the relationship between hourly wages and education mismatch, we use years of education as a proxy for educational attainment and job requirement. This enables us to compare our results with other empirical studies which also assess the relationship between wages and education mismatch using years of education measure.

A range for the required level of education for a particular occupation is established by calculating the mean level of highest educational attainment within each three-digit SOC occupation group. The range is defined as being one standard deviation above and below the mean level of educational attainment. Each individual is then assigned a status based on whether their own level of education falls within or outside of this range for their particular occupation. Table 1 in our previous article gives an illustrated example of this. This method will, by construction, always result in a proportion of workers who can be classified as either:

- matched (individuals whose highest qualification falls within one standard deviation of the average level of educational attainment for their occupation)
- overeducated (individuals whose highest qualification is above one standard deviation of the average level of educational attainment for their occupation)
- undereducated (individuals whose highest qualification is below one standard deviation of the average level of educational attainment for their occupation)

All employees are compared to the average level of educational attainment for the occupation they are in. Aggregating these groups over all occupations gives an estimated matched, over and undereducated rate for the whole economy.

Our adopted methodology differs from the methodology applied in our previously published article as we incorporate several methodological changes and robustness checks. First, we use the Annual Population Survey (APS), whereas previous analysis uses the Labour Force Survey (LFS). Compared to the LFS, the APS provides a larger sample which increases the robustness of our results and enables us to conduct more granular analysis by region and type of degree.

Second, we should emphasise that the statistical method, by its construction, permits the average job requirement to increase across all occupations if participation in education and the average level of educational attainment in the population increases. The effect on the degree of matching across the whole economy is therefore dependent on the age composition of each occupational group and the distribution of older and younger workers across occupations.

For example, as older people leave the labour market – other things being equal – this will tend to increase the average job requirement level for the whole economy, and reduce the percentage of the labour market that are classified as undereducated. To address this issue we construct estimates of required education by three-level SOC separately for two age groups (i) 16 to 35 years and (ii) 36 to 64 years. However, we do acknowledge that this approach may not entirely eliminate the age driven cohort effect as our age bands are somewhat arbitrary.

Third, as a robustness check we construct four different estimates of education mismatch and compare these to an experts view (we use O*NET as a benchmark). We present the comparison between statistical measures and O*NET in Figure 7.

We construct four measures of education attainment using the LFS. Table 5 describes each of these four measures, their advantages and limitations.

Table 5: Statistical measures of education mismatch

Measure	Description	Advantages	Disadvantages
1.Mean level of educational attainment	The mean level of educational attainment is computed for each occupation at each given time period	Relatively easily measurable Changes over time to remain current and up to date	Susceptible to cohort bias
2. Mode level of educational attainment	The mode of the highest level of qualification is computed for each occupation at each given time period. In cases where there are two modes, we take the lower educational level	Less susceptible to cohort bias by taking the most frequent level of education	Not a continuous variable, can only take finite values Volatile time series
3.Mean years of education	Years of education is estimated by subtracting five years from the individuals' continuous years of full time education. The mean is then computed for each occupation at each given time period	Continuous variable measured in years Changes over time to remain current and up to date	Will not capture the true years of education for those with a gap year in their education, nor will it capture the true years of education for those in part time education. Refers to years rather than type /level or other characteristics of education
4.Mode years of education	The mode years of education is computed for each occupation at each given time period. In cases where there are two modes, we take the lower value for the mode years of education	Less susceptible to cohort biases by taking the most frequent years of education	Difficult to establish a suitable threshold to classify matched individuals Will not capture the true years of education for those with a gap year in their education, nor will it capture the true years of education for those in part time education.

Source: Office for National Statistics

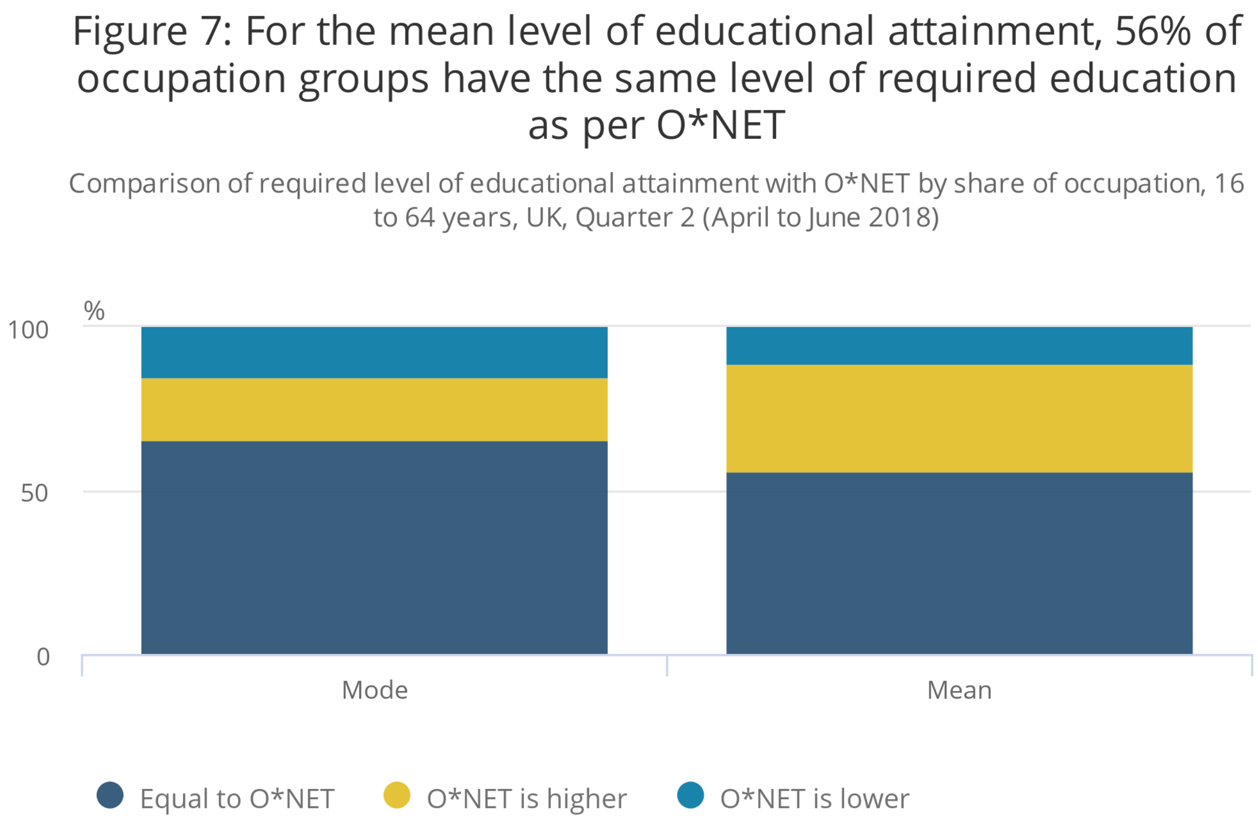
The O*NET category assignment data represents experts' insights into the required level of education for each occupation and it is not subject to cohort bias in the same way that our statistical measures are. Hence, O*NET provides a useful benchmark against which to compare our statistical measures. To conduct this comparison, we map each US SOC to a corresponding four-digit UK SOC, using a comprehensive US to UK SOC lookup.

After mapping all occupations to the LFS for the period April to June 2018, we compare the required level of education in O*NET to our measures (mean and the mode level of educational attainment for each occupation). The educational attainment data we drew from LFS do not always closely reflect the education assignment in the category system as depicted by experts view via O*NET. One major difference is that the category system such as O*NET reflects typical entry-level educational requirements, whereas LFS and APS data report the level of education attained by workers already in the occupation.

We thus observe that for the mode level of educational attainment approximately 66% of occupation groups on a four-level SOC have the same required level of education as per O*NET, whereas this was slightly lower for the mean level of education at 56% (Figure 7).

Figure 7: For the mean level of educational attainment, 56% of occupation groups have the same level of required education as per O*NET

Comparison of required level of educational attainment with O*NET by share of occupation, 16 to 64 years, UK, Quarter 2 (April to June 2018)



Source: Labour Force Survey – Office for National Statistics

Figure 7 also illustrates that there is a considerable share of occupations where category assignment (as per O*NET) shows higher level of education compared with our LFS mean and mode measures. This is slightly higher for the mean, whereby 22% of occupations require a higher level of education as per O*NET, compared with 19% for the mode. It is worth noting that US and UK qualifications are not directly comparable, hence, we have made slight adjustments to account for the fact that students finish high school at the age of 18 years under the US system.

There are several reasons why the educational attainment data may not match the category assignment. Examples are: underemployment, individual choice, and the trend of “upskilling,” in which the educational attainment of workers continues to rise over time. In some cases, the category assignment reflects a higher level of education than the attainment data show. This is because of changing entry requirements, individuals entering an occupation may need a higher level of formal education than those persons who are already working in it.

For example, automotive service technicians and mechanics entering the occupation may need a higher level of education compared to incumbents. Today's automotive engines and components have greater electronic and mechanical complexity, and prospective automotive service technicians and mechanics learn how to repair and maintain them while completing postsecondary education programs. The typical entry level education needed for this occupation is higher than in the past. By contrast, the educational attainment data are only a picture of the recent workforce and may not reflect typical requirements for new entrants to the occupation.

Despite the mode being reported as closer to the experts' view, in our descriptive analysis we chose to use the mean level of education attainment measure whilst making adjustments to mitigate the cohort bias as explained previously. This is because using the mode to produce aggregate matched and undereducated rates poses some limitations. First, with the mode measure we are unable to apply the standard deviation as a threshold from the mode. This would result in significantly lower matched rates of around 40%, therefore implying that 60% of workers are mismatched (either undereducated or overeducated).

Second, using the mode level of educational attainment results in a volatile time series. Given that the mode level of education is finite, a change in the mode would be staggered. For example, from 2012 to 2013 the mode level of educational attainment for a given occupation, may change from 5.0 to 4.0 which will result in volatile changes to the matched and mismatched rates making it more difficult to compare these rates over time.

In comparison with the mean where the level of educational attainment may change from, for example 5.4 to 4.8, the corresponding standard deviation change slightly, resulting in subtle changes in matched and mismatched rates. Therefore, using the mean level of educational attainment with a cohort adjustment allows us to construct a comparable time series while also addressing the cohort bias.

10 . Appendix 2

Regression analysis

In our regression analysis, we control for various individual and job characteristics which may impact both the incidence of overeducation as well as wages. These are: age and age squared (as the relationship between wages and age is non-linear), sex, marital status, number of dependent children, ethnicity, disability, type of employment contract (permanent versus temporary), full time or part time employment, region, occupational skill level, employer size, proxy for social status, immigration status, job tenure (a proxy for organisation specific experience) and survey proxy answers.

Individual factors such as gender, marital status and number of children, are likely to be important determinants of both hourly earnings and educational mismatch. For example, decisions regarding the family may affect individuals' chances of finding work related to their field of study, and this will affect their wages. Staying at home to raise children may require an individual to seek employment with a more flexible schedule, and this may involve foregoing the chance to work in a better paid job or indeed a job for which their qualification was intended.

The descriptive analysis in section three has shown a large regional variation in educational mismatch, implying that both the supply of and demand for skills differ across the UK regions and countries. The economic health and industrial composition of the region or country may largely determine the demand for particular skills and subsequently wages. Also, there is evidence that for certain groups such as married women, labour markets tend to be more geographically restricted. Hence the region or country is another important factor that we account for in our analysis.

Social status can also affect both wage levels and education mismatch (Forest, 2015), and it is usually proxied by the level of education of the father or mother. However, this information is missing in our dataset. We therefore make use of data on home ownership as related evidence shows that this is positively correlated with social status (Forest, 2015; Dang and others, 2017).

We also use the following interaction terms to reflect where the impact of one determinant of wages is affected by the level of another variable: sex multiplied by age, sex multiplied by marital status and sex multiplied by the number of dependent children. Here we acknowledge that females may experience more career interruptions than males due to marriage and having dependent children. Table 6 below is long version of Table 1. Table 6 presents coefficient estimates based on the two specifications of the wage equation discussed in section four with a full set of results for our individual and job control variables. The base control categories in Models 1 to 6 in Table 6 are: rented property; low job skill; London and agriculture.

Table 6: OLS Regression, hourly wage estimation results, 2017 (full model)

Variables	Model 1 log_hourpay	Model 2 log_hourpay	Model 3 log_hourpay	Model 4 log_hourpay	Model 5 Female log_hourpay	Model 6 Male log_hourpay
Obtained education	0.084***	0.076***				
	(0.003)	(0.003)				
Obtained education squared	-0.002***	-0.001***				
	(0.000)	(0.000)				
Overeducated dummy (mean)	-0.081***					
	(0.008)					
Undereducated dummy (mean)	0.058***					
	(0.008)					
Age	0.041***	0.039***	0.049***	0.042***	0.044***	0.050***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of dependent children	0.008**	0.008***	0.010***	0.009***	-0.006*	0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Married	0.072***	0.073***	0.068***	0.072***	0.017***	0.060***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.007)
Female	0.027*	0.027*	0.016	0.026*		
	(0.015)	(0.015)	(0.015)	(0.015)		
White	0.054***	0.054***	0.049***	0.053***	0.023**	0.075***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.011)
UK born	0.058***	0.059***	0.022***	0.036***	0.033***	0.009
	(0.007)	(0.007)	(0.006)	(0.007)	(0.009)	(0.009)
Disability	-0.046***	-0.046***	-0.049***	-0.050***	-0.041***	-0.058***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.009)
Urban location	0.020*	0.021*	0.018	0.020*	0.037**	-0.002
	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)	(0.018)
Private sector	0.003	0.002	0.013**	0.003	-0.006	0.032***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)
Full time	0.051***	0.052***	0.049***	0.052***	0.039***	0.089***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.011)
Permanent	0.063***	0.063***	0.065***	0.057***	0.035***	0.092***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.019)
Small firm	-0.131***	-0.132***	-0.128***	-0.134***	-0.092***	-0.165***

	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)
Proxy answer	-0.034***	-0.034***	-0.035***	-0.038***	-0.011*	-0.052***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.006)	(0.006)
Property owned outright	0.068***	0.069***	0.074***	0.081***	0.056***	0.090***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)
Property bought with mortgage	0.132***	0.133***	0.134***	0.142***	0.108***	0.159***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
Age*Female	-0.001***	-0.001***	-0.002***	-0.002***		
	(0.000)	(0.000)	(0.000)	(0.000)		
Married*Female	-0.059***	-0.060***	-0.057***	-0.057***		
	(0.009)	(0.009)	(0.009)	(0.009)		
Dependent children*Female	-0.016***	-0.015***	-0.015***	-0.015***		
	(0.004)	(0.004)	(0.004)	(0.004)		
High job skill	0.522***	0.542***	0.374***	0.551***	0.416***	0.353***
	(0.008)	(0.008)	(0.010)	(0.008)	(0.015)	(0.014)
Uppermiddle job skill	0.323***	0.333***	0.264***	0.348***	0.283***	0.262***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.011)	(0.010)
Lower middle job skill	0.119***	0.123***	0.104***	0.129***	0.099***	0.115***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)
North East	-0.239***	-0.239***	-0.247***	-0.250***	-0.228***	-0.264***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.014)
North West	-0.222***	-0.222***	-0.229***	-0.229***	-0.218***	-0.240***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.013)
Merseyside	-0.197***	-0.196***	-0.208***	-0.209***	-0.202***	-0.214***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)	(0.017)
Yorkshire and Humber	-0.229***	-0.229***	-0.236***	-0.238***	-0.219***	-0.253***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)
East Midlands	-0.230***	-0.230***	-0.236***	-0.239***	-0.212***	-0.262***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.013)	(0.015)
West Midlands	-0.206***	-0.206***	-0.211***	-0.213***	-0.207***	-0.215***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.013)
East	-0.149***	-0.149***	-0.154***	-0.156***	-0.162***	-0.148***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.013)	(0.015)
South East	-0.120***	-0.120***	-0.125***	-0.124***	-0.135***	-0.115***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)
South West	-0.201***	-0.200***	-0.202***	-0.204***	-0.192***	-0.213***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.013)
West	-0.237***	-0.237***	-0.238***	-0.242***	-0.224***	-0.253***

	(0.008)	(0.008)	(0.008)	(0.009)	(0.011)	(0.013)
Scotland	-0.195***	-0.195***	-0.197***	-0.203***	-0.180***	-0.216***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.013)	(0.014)
Northern Ireland	-0.252***	-0.251***	-0.248***	-0.254***	-0.226***	-0.272***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.017)	(0.018)
Tenure	0.008***	0.008***	0.008***	0.008***	0.009***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Tenure squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Manufacturing	0.099***	0.099***	0.100***	0.102***	0.115***	0.080***
	(0.025)	(0.025)	(0.024)	(0.024)	(0.039)	(0.031)
Energy	0.215***	0.215***	0.212***	0.217***	0.223***	0.196***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.045)	(0.033)
Construction	0.121***	0.122***	0.137***	0.121***	0.157***	0.126***
	(0.026)	(0.026)	(0.025)	(0.025)	(0.042)	(0.032)
Hotels and restaurants	-0.025	-0.024	-0.025	-0.025	-0.030	-0.034
	(0.024)	(0.024)	(0.024)	(0.024)	(0.038)	(0.030)
Transport	0.167***	0.170***	0.160***	0.173***	0.180***	0.137***
	(0.025)	(0.025)	(0.024)	(0.025)	(0.039)	(0.031)
Finance	0.187***	0.192***	0.171***	0.200***	0.167***	0.155***
	(0.025)	(0.025)	(0.024)	(0.024)	(0.038)	(0.031)
Public services	0.017	0.020	0.002	0.026	-0.006	-0.000
	(0.024)	(0.024)	(0.024)	(0.024)	(0.038)	(0.031)
Other services	0.015	0.019	0.011	0.028	0.036	-0.035
	(0.026)	(0.026)	(0.026)	(0.026)	(0.039)	(0.034)
Overeducated dummy (mode)		-0.033***				
		(0.007)				
Undereducated dummy (mode)		0.024***				
		(0.009)				
Required education years (mean)			0.100***		0.082***	0.111***
			(0.003)		(0.004)	(0.004)
Overeducation years (mean)			0.013***		0.010***	0.017***
			(0.002)		(0.002)	(0.002)
Undereducation years (mean)			-0.100***		-0.106***	-0.089***
			(0.006)		(0.008)	(0.009)

Required education years (mode)				0.031***		
				(0.001)		
Overeducation tears (mode)				0.026***		
				(0.001)		
Undereducation years (mode)				-0.082***		
				(0.005)		
Constant	0.289***	0.396***	-0.299***	0.724***	0.058	-0.552***
	(0.051)	(0.051)	(0.057)	(0.047)	(0.082)	(0.081)
Observations	76,360	76,360	76,337	76,337	40,342	35,995
R-squared	0.432	0.430	0.433	0.424	0.412	0.426

Source: Annual Population Survey – Office for National Statistics

Notes

1. Standard errors in parentheses. Back to table
2. *** p<0.01, ** p<0.05, * p<0.1. Back to table

Table 7: OLS regression, hourly wage estimation results for graduates, 2017 (full model)

Variables	Model 7 log_hourpay	Model 8 log_hourpay	Model 9 log_hourpay	Model 10 log_hourpay	Model 11 log_hourpay
Required education	0.054***	0.054***	0.054***	0.053***	0.054***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Overeducation	-0.014***	-0.018***	-0.022***	-0.018***	-0.017***
	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
Recent graduate	-0.005	-0.003	-0.005	-0.016	-0.005
	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)
STEM degree subject	0.021**	0.021**	0.021**	0.021**	0.016*
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Age	0.064***	0.065***	0.064***	0.063***	0.064***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of dependent children	0.017**	0.017**	0.018**	0.018***	0.017**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Married	0.083***	0.084***	0.082***	0.082***	0.083***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Female	0.018	0.017	0.014	0.019	0.017
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
White	0.055***	0.055***	0.054***	0.055***	0.055***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
UK born	0.078***	0.077***	0.078***	0.077***	0.077***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Disability	-0.052***	-0.052***	-0.052***	-0.053***	-0.052***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Urban location	0.049**	0.048*	0.050**	0.050**	0.048*
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Private sector	0.047***	0.047***	0.047***	0.047***	0.048***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Full time	0.064***	0.064***	0.064***	0.064***	0.064***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Permanent	0.048**	0.048**	0.047**	0.047**	0.048**
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Small firm	-0.168***	-0.167***	-0.168***	-0.168***	-0.168***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Proxy answer	-0.018**	-0.018**	-0.019**	-0.018**	-0.018**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)

Property owned outright	0.046***	0.046***	0.046***	0.046***	0.046***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Property bought with mortgage	0.102***	0.102***	0.102***	0.102***	0.102***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Age*Female	-0.001*	-0.001*	-0.002**	-0.001*	-0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Married*Female	-0.056***	-0.056***	-0.054***	-0.056***	-0.056***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Dependent children*Female	-0.032***	-0.032***	-0.033***	-0.033***	-0.032***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
High job skill	0.401***	0.399***	0.399***	0.401***	0.398***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Uppermiddle job skill	0.315***	0.313***	0.314***	0.315***	0.312***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Lower middle job skill	0.137***	0.136***	0.135***	0.138***	0.136***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
North East	-0.278***	-0.278***	-0.278***	-0.278***	-0.278***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
North West	-0.273***	-0.273***	-0.272***	-0.273***	-0.273***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Merseyside	-0.229***	-0.229***	-0.229***	-0.229***	-0.229***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Yorkshire and Humber	-0.268***	-0.268***	-0.268***	-0.268***	-0.268***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
East Midlands	-0.267***	-0.267***	-0.267***	-0.267***	-0.267***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
West Midlands	-0.246***	-0.246***	-0.246***	-0.246***	-0.246***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
East	-0.161***	-0.161***	-0.161***	-0.161***	-0.161***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
South East	-0.147***	-0.147***	-0.147***	-0.147***	-0.147***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
South West	-0.256***	-0.256***	-0.255***	-0.255***	-0.255***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
West	-0.263***	-0.262***	-0.262***	-0.262***	-0.262***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)

Scotland	-0.244***	-0.243***	-0.244***	-0.244***	-0.244***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Northern Ireland	-0.284***	-0.284***	-0.283***	-0.283***	-0.284***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Tenure	0.010***	0.010***	0.010***	0.010***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Tenure squared	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Manufacturing	0.160**	0.159**	0.159**	0.160***	0.160**
	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
Energy	0.258***	0.257***	0.257***	0.259***	0.259***
	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)
Construction	0.145**	0.144**	0.144**	0.146**	0.146**
	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)
Hotels and restaurants	-0.002	-0.002	-0.001	-0.001	-0.001
	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
Transport	0.154**	0.153**	0.154**	0.155**	0.154**
	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
Finance	0.188***	0.187***	0.187***	0.189***	0.188***
	(0.061)	(0.062)	(0.062)	(0.061)	(0.061)
Public services	0.007	0.006	0.007	0.008	0.007
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Other services	-0.012	-0.013	-0.012	-0.011	-0.011
	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)
Overeducation*Tenure		0.000*			
		(0.000)			
Overeducation*Female			0.014***		
			(0.005)		
Overeducation*Recent graduate				0.012**	
				(0.005)	
Overeducation*STEM degree subject					0.011*
					(0.005)
Constant	-0.024	-0.027	-0.010	0.008	-0.024
	(0.141)	(0.141)	(0.141)	(0.142)	(0.141)
Observations	17,672	17,672	17,672	17,672	17,672
R-squared	0.403	0.403	0.403	0.403	0.403

Source: Annual Population Survey – Office for National Statistics

Notes

1. Robust standard errors in parentheses. Back to table

2. *** p<0.01, ** p<0.05, * p<0.1. Back to table

Table 8: Quantile regression, hourly wage estimation results for graduates, 2017 (full model)

	Model 12	Model 13	Model 14
Variables	**q25**	**q50**	**q75**
Required education	0.040***	0.045***	0.061***
	(0.005)	(0.004)	(0.004)
Overeducation	-0.010***	-0.011***	-0.011***
	(0.002)	(0.003)	(0.003)
Recent graduate	-0.009	-0.013	-0.004
	(0.013)	(0.012)	(0.016)
STEM degree subject	0.016*	0.011**	0.007
	(0.008)	(0.005)	(0.006)
Age	0.047***	0.052***	0.065***
	(0.002)	(0.003)	(0.004)
Age squared	-0.000***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Number of dependent children	0.016***	0.016**	0.029***
	(0.005)	(0.007)	(0.010)
Married	0.068***	0.076***	0.075***
	(0.014)	(0.014)	(0.017)
Female	0.047*	0.055***	0.108***
	(0.025)	(0.021)	(0.026)
White	0.042***	0.044***	0.061***
	(0.011)	(0.011)	(0.018)
UK born	0.080***	0.073***	0.067***
	(0.012)	(0.013)	(0.017)
Disability	-0.038***	-0.048***	-0.073***
	(0.011)	(0.009)	(0.011)
Urban location	0.021	0.028	0.024
	(0.020)	(0.021)	(0.021)
Private sector	0.011	0.046***	0.067***
	(0.010)	(0.009)	(0.013)
Full time	0.068***	0.063***	0.061***
	(0.009)	(0.008)	(0.013)
Permanent	0.062***	0.030**	-0.024
	(0.017)	(0.015)	(0.017)
Small firm	-0.118***	-0.107***	-0.105***
	(0.008)	(0.008)	(0.011)
Proxy answer	-0.015**	-0.025***	-0.004
	(0.007)	(0.007)	(0.011)

Property owned outright	0.037***	0.066***	0.086***
	(0.011)	(0.008)	(0.011)
Property bought with mortgage	0.099***	0.111***	0.122***
	(0.010)	(0.007)	(0.008)
Age*Female	-0.002***	-0.003***	-0.004***
	(0.000)	(0.001)	(0.001)
Married*Female	-0.033	-0.039**	-0.042*
	(0.021)	(0.017)	(0.024)
Dependent children*Female	-0.034***	-0.020*	-0.019
	(0.009)	(0.011)	(0.012)
High job skill	0.421***	0.460***	0.480***
	(0.020)	(0.017)	(0.019)
Uppermiddle job skill	0.299***	0.341***	0.388***
	(0.013)	(0.018)	(0.020)
Lower middle job skill	0.115***	0.124***	0.175***
	(0.013)	(0.018)	(0.020)
North East	-0.260***	-0.286***	-0.329***
	(0.016)	(0.018)	(0.020)
North West	-0.247***	-0.274***	-0.308***
	(0.014)	(0.016)	(0.020)
Merseyside	-0.226***	-0.223***	-0.304***
	(0.020)	(0.024)	(0.020)
Yorkshire and Humber	-0.264***	-0.268***	-0.318***
	(0.018)	(0.018)	(0.024)
East Midlands	-0.250***	-0.259***	-0.304***
	(0.011)	(0.015)	(0.021)
West Midlands	-0.235***	-0.233***	-0.303***
	(0.015)	(0.015)	(0.017)
East	-0.205***	-0.188***	-0.188***
	(0.015)	(0.022)	(0.027)
South East	-0.157***	-0.152***	-0.164***
	(0.012)	(0.016)	(0.017)
South West	-0.231***	-0.253***	-0.289***
	(0.018)	(0.016)	(0.022)
West	-0.250***	-0.274***	-0.321***
	(0.012)	(0.017)	(0.018)
Scotland	-0.211***	-0.234***	-0.284***
	(0.013)	(0.021)	(0.019)

Northern Ireland	-0.246***	-0.283***	-0.357***
	(0.030)	(0.022)	(0.023)
Tenure	0.013***	0.012***	0.008***
	(0.001)	(0.001)	(0.001)
Tenure squared	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Manufacturing	0.199***	0.105	0.036
	(0.056)	(0.073)	(0.056)
Energy	0.290***	0.202***	0.156***
	(0.060)	(0.076)	(0.051)
Construction	0.176***	0.091	0.003
	(0.061)	(0.070)	(0.053)
Hotels and restaurants	0.016	-0.074	-0.116**
	(0.059)	(0.071)	(0.054)
Transport	0.189***	0.097	0.039
	(0.064)	(0.063)	(0.054)
Finance	0.197***	0.117*	0.067
	(0.057)	(0.070)	(0.054)
Public services	0.068	-0.049	-0.151***
	(0.059)	(0.068)	(0.050)
Other services	0.018	-0.046	-0.100*
	(0.066)	(0.068)	(0.055)
Constant	0.291**	0.392***	0.156
	(0.124)	(0.123)	(0.117)
Observations	17,672	17,672	17,672

Source: Annual Population Survey – Office for National Statistics

Notes

1. Standard errors in parentheses. Back to table
2. *** p<0.01, ** p<0.05, * p<0.1. Back to table