

Real Wages, Wage Inequality and the Regional Cost-of-living in the UK

Cinzia Rienzo*

First Draft November 2010

Abstract

University graduates in the UK are more concentrated in regions where the cost of housing is higher than other groups. This implies that many graduates face a higher cost-of-living that could possibly reduce the graduate real wage relative to other groups and carry implications for measures of wage dispersion.

This paper aims to reassess how estimates of wage inequality from 1997 to 2008 vary when regional differences in the cost of housing in the UK are taken into consideration. In order to do so, the real wage is deflated by a specially constructed regional Retail Price Index (RPI); this is a new measure of the cost-of-living that partially updates the national RPI with a regional housing index, therefore allowing the RPI to vary by regions.

Results show that the national RPI underestimates the cost-of-living of workers living in the most expensive regions (London, South East) and overestimates the cost-of-living for “cheaper” regions (Northern Ireland, Scotland). When deflating hourly wages by the regional RPI, the average level of wages is lower from 8% to 11% an hour for all workers in London and the South East, but is higher (from 2% to 9%) in the remaining regions; similarly wage inequality is lower when deflating by the real regional RPI.

Key words: cost-of-living, wages, wage inequality

JEL Classification: E3, J31

* Royal Holloway, University of London and Centre of Migration Policy and Society *COMPAS*, University of Oxford; e-mail c.rienzo@rhul.ac.uk.

I am grateful to Jonathan Wadsworth for several suggestions and encouragement. I also would like to thank Ian Crawford and Jim O’ Donoghue for fruitful discussions. My thanks to Enrico Moretti for providing me with his do file. I am also grateful to seminar participants at the RES Annual Conference 2011 and WPEG Annual Conference 2011 for comments and suggestions. Remaining errors are my own. The Expenditure and Food Survey, Labour Force Survey and the Family Resources Survey data are made available through the UK Data Archive (UKDA).

1. Introduction

The estimated wage returns to graduates has increased significantly in the UK over time (Card and Lemieux, 2001; Machin, 1998; Machin and Van Reenen, 2008; Rienzo, 2010; Schmitt, 1995). Beginning in the 1980s, despite a continued expansion in the relative supply of skilled workers, rising relative demand for skilled workers led to widening skills differentials (Schmitt, 1995); Card and Lemieux (2001) claim that the increase in the college high school wage gap for the US, UK and Canada over the past 30 years is attributable to steadily rising relative demand for college-educated labour but in different sub-periods the slowdown in graduate returns is related to rising relative supply. Machin (2003) documents that over the 1980s the returns to graduate education in the UK were rising and in the 1990s were relatively stable. Recent studies also demonstrate that the average returns to graduate education remain high (Sloane, 2003; Walker and Zhu, 2003). Green and Zhu (2010) also find that the dispersion of the returns to graduate education substantially increased for both men and women over the period 1994 to 2006.

In a recent contribution Moretti (2010) questioned the relative real wage increases for US graduates by re-examining how wage inequality is measured. He demonstrates how existing estimates of wage inequality for the US change when accounting for differences in the cost-of-living across locations and the relative concentration of graduates in certain high cost locations. To do so he deflates nominal wages using a new CPI that allows the cost of housing to vary across metropolitan areas. Using data from the US Census between 1980 and 2000, Moretti focuses on the difference in the average hourly wage for workers with a high school degree and workers with college degree or more. Much of the growth in the number of college graduates has occurred in metropolitan areas that have both a higher initial cost of housing and which have experienced larger increases in the cost of housing over time. This implies that college graduates are increasingly exposed to a higher cost-of-living and that the relative increase in their real wage may be smaller than the relative increase in their nominal wage. He shows that between 1980 and 2000, the cost of housing for college graduates grew much faster than the cost of housing for high school graduates. In 1980 the difference in the average cost of housing between college and high school graduates was only 4%. The difference grew to 14% in 2000. Moreover this implies that the difference between the wages of college graduates and high school graduates is smaller in real terms than in nominal terms. In fact half of the documented increase in the return to college between 1980 and 2000 disappears when using real local CPI to deflate wages.

This paper attempts to address this issue for the UK where similar trends in wage inequality, changes in educational characteristics of workers, and concentration of graduate workers in more expensive regions are observed. Card and Lemieux (2001), Gosling and Lemieux (2004), Machin

and Van Reenen (2008) and Rienzo (2010) highlight the similarities between the UK and the US in both the increase in wage inequality and changes in educational characteristics of the labour force. The UK, like the US, over the past four decades experienced a secular increase in both the 90-50 and 90-10 wage differential (Bell and Van Reenen, 2010). Both countries experienced increases in the proportion of graduate workers in the population. In the US this grew from 20.8% in 1980 to 34.2% in 2004; for the UK the growth in graduates was from 5% to 21% over the same time period. Based on the Labour Force Survey (LFS) Leunig and Overman (2008) show that the proportion of graduate households with two earners increased from 53% in 1977 to 77% in 2006.

As in the US, graduate workers in the UK are also unevenly distributed across the country. Looking at the distribution of graduates across the UK, based on the LFS in 2008, nearly 38% of the UK graduate workers were concentrated in two regions: the London area (21.1%) and the South East (16.7%); those two regions also made up respectively 11.5% and 15% of the total UK labour force in 2008. In London, only 6.2% of workers have less than a high school level of education and 11.4% have a high school degree; in the South East about 29.2% of workers have a high school degree (16.5%) and less than high school (13.1%). These are also the British regions where households spend more on housing than the UK average, largely due to differences between regions in the average amount spent on rent and mortgages (Family Spending, ONS 2009). Since the late 1970s, there have been considerable variations in house prices both over time and across UK regions,¹ with prices in London rising faster than other regions (Holly et al. 2010). Similarly between 1997 and 2008, graduate workers became more concentrated in London. Therefore it is likely that changes in house prices may affect real wage levels of graduates more over time; and this is another reason motivating this study.

The distribution of graduate workers raises the question as to why graduate workers tend to concentrate into more expensive areas. Explanations can be found in the relative demand and relative supply of skilled workers (Moretti, 2010). Because firms in more expensive cities experience an increase in the relative demand for skilled workers, more educated workers move to those more expensive cities. Therefore the increase in their utility level is smaller than their increase in their nominal wage. In this scenario the increase in well-being inequality is smaller than the increase in nominal wage inequality because of their higher cost-of-living. The other explanation related to the concentration of graduate workers in more expensive cities can be found in the increase of relative supply due to an increase of local amenities that attract graduate workers;

¹ For example from January 1999 to July 2004 house price in the UK rose by 123% (Nickell, 2005)

Moretti (2010) explains that the increase in the cost-of-living is somehow correlated with the increased attractiveness of those cities.

In a similar vein to Moretti (2010), Black et al (2010) question whether the return to education is likely to be the same in locations characterised by differential price levels. They demonstrate that in an equilibrium that has local variation in prices, not only do wage levels differ across locations, but so too do returns to schooling. They examine the returns to college education relative to high school education for large cities in the US in 1980, 1990 and 2000 and find persistent and substantial heterogeneity in the return to a college degree, supporting also the prediction that the local return to schooling is inversely related to housing prices.

The existing literature on wage inequality does not reach any consensus on what caused the secular rise in wage dispersion both within and across countries. Researchers do agree that the causes seem to lie with a variety of components, rather than one exclusive factor, that jointly affect the wage structure. In particular, the leading explanations regarding rising wage inequality such as declining unionization, the falling real value of the minimum wage, increased trade and skill-biased technological change do not seem to help explain recent trends in wage inequality since they have less to say regarding the dominant trends of the 1990s and 2000s, namely increasing upper-tail inequality and declining lower tail inequality (Autor, Katz and Kearney 2005). These factors also fail to explain that residual wage dispersion among workers with the same education and experience accounts for most of the growth in overall wage inequality (Lemieux 2006).

The existing literature investigating trends and causes of wage inequality in the UK usually measures wages in real terms by deflating nominal wages using the national Retail Prices Index (RPI). The RPI provided by the Office for National Statistics (ONS) is a fixed quantity price index: it measures the proportional change in the cost of buying some fixed bundle of goods as prices change (Blow and Crawford, 2001). However, the RPI does not account for differences in regional housing costs.

Expenditure on housing represents the largest component of total household expenditure but this varies considerably across regions in the UK. As such, differences in regional housing costs might be expected to play an important part in determining cost-of-living differences between regions. For example in 2008, housing costs accounted for 34% of total expenditures of people renting in London compared to 24% for those living in Wales and 18% for inhabitants of Northern Ireland (Family Expenditure Survey (FES), 2008). In the same year, in London 43% of the labour force population were graduates, compared to 21% in Wales and 24% in Northern Ireland (LFS). So depending on where they live, workers with the same level of education may face a different cost-of-living. This implies that deflating the nominal wage by a regional RPI might lead to different

estimates of the observed real wage dispersion than one based on a national index that might fail in being fully representative at the regional level.

During the last two decades a few studies addressed how accounting for regional variation of prices in the UK could reflect in different estimates of income or inequalities, though none of them investigated how this could affect wage dispersion and return to education.

For example Crawford (1996) examined the extent and pattern of differences in the cost-of-living for subgroups of the population, specifically income group and head's date-of-birth cohort. He pointed out that to a large extent the difference in cost-of-living inflation for households in the top and bottom 10 per cent of income distribution is driven by variations in housing tenure types between the two groups. He demonstrates that adding housing costs increases the average difference in inflation rates for poorer households relative to richer households from -0.01 to -0.07 percentage points. The same author also provides evidence of cohort differences in inflation rates between 1979 and 1992 for three broad head of household birth cohorts: households in which the head was born before 1930, those in which the head was born after 1930 but before 1960, and those in which the head was born after 1960. For those born after 1960, their cost-of-living at the end of the period had grown 2.68 per cent more than average; in contrast to them the eldest households did relatively well, finishing the period with a cost-of-living which had grown 0.45 per cent slower than average.

Borooah et al. (1996) used regional price data provided by the Croner Reward Group (CRG) in conjunction with weights derived from the FES to construct regional retail price indices. The relative expensiveness of Greater London and the South East, over the period 1979-1990, increased when housing costs were included; conversely, the inclusion of housing costs meant that Northern Ireland changed from being slightly more, to slightly less expensive than the UK average. They found that Northern Ireland was above average in the cost of: food, housing repairs, fuel and travel, and below average in the cost of leisure services. Greater London had relatively high costs for: food, fuel, personal goods and services. When housing costs were included in the aggregate index, Greater London and the South East increased their mean relative cost index value to 1.052 and 1.020, from 1.023 and 1.005 respectively. On the other hand, the inclusion of housing led Northern Ireland to experience a fall in relative costs of a magnitude comparable to London's increase.

Acknowledging the high degree of expenditure variation rates amongst different households in the UK, Crawford and Smith (2002) stipulate a household specific cost-of-living indices reflecting that households consume goods and services in different proportions. Using data from the UK FES from 1975 to 1999, they analyse differences in inflation rates by grouping households according to whether they pay rent for their accommodation, whether they own it outright or whether they own it with a mortgage. They use the Laspeyres index to generate household specific

inflation; the index compares the cost of buying the observed set of commodities at two different sets of prices: the contemporaneous and the set prevailing in the following period. The differences in the inflation rates for the households are generated by differences in their commodity demands.

Crawford and Smith (2002) claim that the headline average inflation rate is not always close to the experience of inflation for “the great majority of households”. They provide evidence that on average from 1976 to 2000 only about one third of households at a point in time faced inflation rates within 1 percentage point of the average rate. Analysing the impact of ignoring differential inflation on the measurement of income inequality, they find that not allowing for differential inflation could lead to the annual growth rate in inequality being overstated or understated by as much as 6 percentage points². The measures of inflation they derive are based on the price data that are in the published section indices of the RPI, collected from national sources that do not account for any regional variation in prices.

This paper differs from the contribution of Crawford and Smith (2002) by accounting for regional variation in the cost-of-living; furthermore while Crawford and Smith (2002) base their analysis on the household disposable income, this paper focuses on the individual hourly wage as a point in time measure of labour, allowing to understand the effect of different measures of inflation on the cost of labour.

Duranton and Monastiritis (2002) investigate regional inequalities in the UK from 1982 to 1997 and their evolution by examining labour market earnings. They document a convergence of wage equalization across UK regions during the time period in both the coefficients on regional fixed effects and in the returns to key labour market characteristics such as experience, education and sex. They show that rising inequalities between skilled and unskilled in combination with the increasing and uneven spatial distribution of education in the UK, contributed to amplify aggregate regional disparities. They also pointed out that the rising average educational attainment in London and the South East relative to the rest of the country played a role in explaining the aggravation of regional inequalities. Their analysis is based on the FES and the General Household Survey (GHS) to derive the log of real weekly earnings and they construct two panel-price indexes based on the

² Crawford and Smith (2002) use a different definition of income than the one used to calculate official low-income statistics. They use household disposable incomes, after adjusting for household size and composition, as a proxy for material living standards. More precisely, it is a proxy for the level of consumption of goods and services that people could attain given the disposable income of the household in which they live. In order to allow comparisons of the living standards of different types of households, income is adjusted to take into account variations in the size and composition of the households in a process known as equivalisation. A key assumption made in HBAI is that all individuals in the household benefit equally from the combined income of the household. This enables the total equivalised income of the household to be used as a proxy for the standard of living of each household member.

regional prices available from the Croner-Reward Group; however they do not construct or use any common price deflator. This paper also aims to do construct a regional common price deflator. Hayes (2005) identifies a number of problems in using these latter data and interpretation of the results based on the construction of the deflator; the main concern being the dominance of house price movements; Hayes points out that if CRG indices are taken as proxies for the published RPI, the indices are not a fair representation of the cost-of-living for households not purchasing or renting a house in each period.

One problem is due to the fact that in the CRG the expenditure total is a measure of a cost-of-living reflecting a lifestyle rather than a collection of prices from a given basket (as used in the RPI). Because of some dissimilarity in the construction of the CRG prices with respect to the RPI, the combination of the two data would not be appropriate. In fact, in constructing the expenditure totals in each region, CRG collects prices data over 100 localities, as well as at national levels, for over 260 specific items. Each year the expenditure items are reviewed and changes are made when necessary. The RPI is published monthly using price data collected on 500-600 specific types of goods and services based on 180 shops throughout the country. Although the methodology to construct the price index in CRG is similar to the one followed by the ONS to construct the RPI, few differences are found; for example unlike the RPI, the CRG includes the top 4% income earnings. Another discrepancy is due to the timing of collection of prices: the RPI is based on data collected each month; the CRG only collects data twice a year. This makes a time series analysis of the behaviour of regional prices problematic.

To address concerns on whether there is local area economic convergence in the UK Henley (2005) uses data on Gross Domestic Product (GDP)/Gross Value Added (GVA) per capita at 128 Nomenclature of Units for Territorial Statistics (NUTS level 3) sub-regions across the UK published by the ONS every year; a sub-region coincides with single or groups of unitary local government authorities. Because local GDP deflators are not available, regional-level analyses of GDP convergence typically do not investigate the impact on convergence estimates of variation in local prices. Henley shows that the use of regional price deflators may affect rates of convergence estimates; in fact in convergence models using country-level data, one finds evidence for positive region-specific convergence; however when accounting for regional differences in the cost-of-living no evidence can be found for the convergence across the the UK.

Despite its potential importance, the existing literature on UK wage inequality has not accounted for differences in the cost-of-living nor analysed whether this could possibly affect the level of and changes in observed wage and wage inequality. The decisions about how we compute inflation statistics can have a direct impact on policy decisions (Checchetti, 2007). The RPI is in

fact used by the government for a number of purposes including the calculation of various incomes and prices; it is used to set and up-rate the level of wages, tax allowances, and to regulate train fares as well as index-linked government bonds. The Government used to refer to the RPI to set welfare benefits and state pensions, however very recently it has been announced? to change the inflation benchmark used to set benefits and pension from the RPI to the Consumer Prices index (CPI) (Inman, P ,The Guardian, 2010). The regional variation in inflation rate is an issue of importance because of the crucial role of the RPI. Because the RPI does not reflect any regional variations, all decisions based on that will not account for the different “real” cost-of-living faced by individuals living in different UK regions. In 2003 the Chancellor of the Exchequer in his budget speech stated that “in future we plan regional price indexes showing differences in regional inflation rates” (Fenwick and O’Donoghue, 2004).

Blanchflower and Oswald (2005) advocate the need for UK regionally-determined wage packages. They argue that because the cost-of-living varies regionally, public sector workers should be paid differently, according to where they work as the level of private sector wages varies dramatically across different parts of the UK; for example, workers in central London earn 55% more than others – so should remuneration in the public sector also take into account regional differences in the cost-of-living(Blanchflower and Oswald, 2003). As pointed out by Borooah et al (1996), if it is true that some regions of the UK are “cheaper” than others and that such differences persist over time, then a policy of paying nationally determined unemployment benefits might have, depending upon the region of residence, significantly different consequences for the real incomes of the unemployed and so affect their participation in the labour market. A study of regional variations in the cost-of-living has several important implications; Borooah et al (1996) pointed out three: first, there is the adjustment of social security benefit levels to take account of regional differences in prices. Secondly, conclusions about the relative deprivation or prosperity of regions, as measured by real disposable income, could also be susceptible to change in the face of regional variations in the cost-of-living. Lastly, conclusions about the number of persons living in poverty could also alter when regional cost-of-living variations are allowed for.

The empirical evidence on wage inequality in the UK is mainly based on the measure of wage inequality derived by deflating nominal wages with the national RPI. Less attention has been paid to regional disparities in wage inequality in the UK and most of the existing measures of the RPI do not account for regional variations in the cost-of-living.

In attempting to shed some light on this area, this chapter focuses on changes in the difference in the log hourly wage for workers with a college degree and workers with a high school degree or workers with a less than high school degree in the UK between 1997 and 2008. Data on

individual wages are based on LFS data; wages are deflated by a regional RPI that is derived by using the national RPI excluding housing expenditure augmented by a weighted price of housing component that varies across 12 regions using weights derived from the FES and price for rents derived from the Family Resources Survey (FRS). The real hourly wage is therefore deflated using a new RPI that allows for variation in the cost-of-living by 12 UK regions. Moreover this paper attempts to investigate the effects of the cost-of-living on the level of wage dispersion over a period in which inequality was static or falling. In fact, from 1997 to 2008 wage inequality in the UK decreased by about 0.021 (Standard deviation) to 0.051 (50-10 gap) log points.

This paper is organised as follows: part two explains the methodology followed by the ONS to construct the RPI and presents some related issues; part two also explains how the regional RPI is derived. Part three presents the econometric methodology; in part four the data sets used are discussed. Part five discusses the results, and part six concludes and discusses future research.

2 RPI and the cost-of-living

This section begins by discussing how the RPI is calculated by the ONS (section 2.1). The second part of this section describes how the Regional RPI used in the main estimates is derived.

2.1 National RPI: descriptions and drawbacks

The RPI is defined (ONS, 2007)³ as an average measure of change in the prices of goods and services bought for the purpose of consumption by the vast majority of households in the UK. The spending pattern underlying the RPI is that of an average private UK based household, excluding certain households: the top 4% of households by income and “pensioner” households where state benefits provide at least 75% of their income. To represent price movement in the RPI basket specific representative items are chosen; there are currently 650 representative items that have fairly broad specifications and collectors must choose a variety that conforms to that specification.

The methodology for deriving the RPI involves weighting together aggregated prices for different categories of goods and services so that each takes its appropriate share within household budgets. The RPI uses aggregate average expenditure to calculate weights implying that each index household contributes to the weights an amount proportional to its expenditure.

³ Consumer Price Indices Technical Manual , page 3, 2007.

The data used to produce the weights⁴ are based on the Expenditure and Food Survey (EFS), renamed the Living Costs and Food Survey (LCF) in 2008 and formerly the FES. The EFS/FES/LCF is a continuous household survey which monitors the spending patterns of around 6,000 to 6,500 household across the country each year.

The RPI is an annually chain-linked index: each year a separate index on the most recent January is produced and each year's indices are then chained together to produce an index covering several years. Within each year the RPI is a fixed quantity price index that measures the change in the price of a basket of fixed composition: a price index such as the RPI measures the proportional change in the cost of buying some fixed bundle of goods as prices change; however as pointed out by Blow and Crawford (2002), when prices change, consumers can mitigate the effect of price increases on their cost-of-living by substituting away from goods that have become relatively more expensive towards those that have become relatively cheaper. Therefore because the RPI ignores the substitution effects it may suffer from substitution bias when compared to a true cost-of-living index. Nevertheless the RPI Technical Manual (ONS, 1998) explains that the RPI is not intended to measure what people often refer to as "the cost-of-living"; a cost-of-living⁵ index measures the average change in prices with reference, not to a fixed list of demands, but to a fixed standard of living (Crawford and Smith, 2002).

Based on the above methodology the ONS provides four different measures of inflation that differ in the exclusion or inclusion of housing costs: the RPI including all items (CHAW); the RPI excluding mortgage interest payments (CHMK); the RPI excluding mortgage interest payments and indirect taxes (CBZW) and the RPI excluding housing (CHAZ).

The RPI all items (CHAW) includes costs of housing at national level. Costs for homeowners are represented by nominal mortgage interest payments (MIPs) that are measured using a model of the payments being paid for mortgagees by an average household; the calculation starts with the average price of new and existing houses bought on mortgages in each year in the past 23 years. Essentially the current approach is to multiply the average outstanding mortgage debt (calculated as a weighted average of the values of mortgages taken out over the previous 23 years) by current interest rates. The main source of house price data is the "mix-adjusted" Department of the Environment, Transport and Regions (DETR) house price index. "Mix-adjusted" means that the sample is weighted to ensure that there is no steep change in the sample average house price due to

⁴ The RPI weights are constructed on the same population basis as the RPI; i.e. excluding the expenditure of the top 4% and pensioner households; these households are excluded because they have very different spending patterns from most households.

⁵ Konus (1924) first defined a true cost-of-living as the minimum cost of achieving some reference welfare level when the price vector is p_t relative to the minimum cost of achieving the same welfare with the price vector p_s .

different sizes or types of properties being recorded each year. This house price index is compiled from sample information on mortgages loaned from building societies, banks and other mortgage lenders. The sample is trimmed specifically for the RPI purposes by removing house transactions where the mortgagors' income given in the mortgage application is over the top 4% threshold used to define index households. The house price estimate is calculated by combining the most recent month's Halifax house price index with earlier months' data from the DETR index. The interest rates used at present are a weighted average of the 20 largest building societies and the nine largest banks' basic mortgage interest rates adjusted. In each month of the 23 year calculations the house price is multiplied by 55% for houses bought after 1981 and by 65% for houses bought before 1981 (these percentages are the assumed proportions of the mortgage that was borrowed). This gives the average current debt outstanding on mortgages for each month of each of the 23 years and those figures are subdivided by eligible tax relief (£30000 or less) and the rest. The two average debt figures are then multiplied by average mortgage interest rates net and gross of tax relief, as charged by building societies. This results in an average mortgage interest payment in each of the two tax categories for owner occupied households. The two indexes are then added together to give the average mortgage interest payments that is multiplied by the proportion of owner-occupiers who have lived in their house for less than 23 years; then by the proportion of under 23 year owner-occupiers with mortgages and then by the proportion of index households which are owner-occupiers. This gives an average payment of all index households that is the measure for housing used in the RPI CHAW.

The RPI CBZW excludes these mortgage interest payments as well as indirect taxes and includes, as housing costs, national rent; information on rents comprising private sector, local authority and social landlord rent. Private sector rents are collected from estate agents in around 150 locations around the country. The price collected is then net of any inclusive water, sewerage or council charges, as these are accounted for by separate centrally collected items. Local authority rent information is provided by the Department of Communities and Local Government (DCLG) and the equivalent bodies within the devolved administrations. Registered social landlord data are from the Tenant Services Authority.

The RPI CBZW also includes council tax and rates; water and other charges; repairs and maintenance charges and dwelling insurance.

The RPI excluding housing (CHAZ) does not include any of the previous housing costs. Any RPI that calculates housing costs based on mortgages is likely to be higher than the RPI that includes rents (CBZW) as a measure for housing costs due to the fact that not only weights for mortgages are larger than those for rents, but also mortgage payments are higher than rents. As

explained by Crawford (1996), these factors together make the cost-of-living indices extremely sensitive to fluctuations in mortgage interest rates; on average a 1 percent increase in mortgage interest rates raises the RPI by 0.5 per cent. In addition, as explained above, the methodology used to calculate the cost for housing based on mortgages differs from the one that uses rents.

However in any of the measures of inflation available from the ONS there are two aspects that the RPI does not account for. The first one is the variation of housing related expenditure by region in the UK, the other one is the difference in the inflation rates that different household types experience across the UK.

This paper focuses on the former. At present the ONS does not calculate regional inflation figures. Fenwick and O'Donoghue (2003) explain why particular features of the RPI sample design and price collection methods are not suitable for regional price level comparisons. First of all the EFS/FES/LCF sample is not designed to be representative of individual regions but it is designed to represent the shopping habits of the average UK consumer and therefore does not reflect differences in the range of items purchased and expenditure shares in different regions of the country. Similarly, the collection of price data is designed for the construction of national price indices only and not regional ones. Regional spending patterns would be required so that regional indices could reflect accurately the consumption basket by residents of the region; the information needed for such an exercise is contained in the ONS's annual Expenditure and Food Survey but the sample size within regions is not sufficient to generate reliable statistics for the regions.

Based on the idea that differences in cost-of-living indices between population groups are generated entirely by their spending pattern and that housing tenure is likely to generate large differences in inflation rates, Crawford and Smith (2002) calculate household-specific cost-of-living indices. They group these indices across households according to whether they pay rent for their accommodation, whether they own it with a mortgage or whether they do not pay any housing costs at all. They show that inflation rates across renters, house owners and mortgage holders display large differences between them. The biggest difference occurs between mortgagors and the other two groups. In 1989, for example, the inflation rate for mortgagors was about 12.9 per cent while for the other two groups it was 6.3 per cent. During the 1990s, mortgagors experienced much more variation in the inflation rate than the renters and owners. This coincides with the increase in mortgage rates during the 1980s. From 1988 onwards, increases in interest rates pushed the cost-of-living of home-owners up faster while rents lagged. The interest rate cuts from the early 1990s had the reverse effect, cutting the rate of increase for home-owners relative to the average while the average rents rose more sharply. The results show that using household-specific price indices gives a different picture of the evolution of inequality in living standards over the period from the one that

emerges if these inflation effects are ignored. For example, between 1978 and 1979 inequality increased on the measure that assumes common inflation rates, whilst the measure that allows for differential inflation effects shows a decline in inequality over the same period.

2.2 Methodology: Constructing a Regional RPI

Throughout this chapter the National RPI refers to the one provided by the ONS (CBZW) excluding mortgage interest payments and indirect taxes and including other national housing costs (gross rents); since the RPI/CBZW is the benchmark against which the derived Regional RPI is compared, both weights and the housing costs are derived by accounting for the housing component of the RPI/CBZW.

The Regional RPI refers to the one that is proposed as a new measure of the cost-of-living taking into account regional differences in housing costs over time and the National RPI/CHAZ is the one that excludes any housing costs and that is used as a base to construct the new measure. Using the RPI/CHAZ as a base is more convenient than using the RPI/CBZW since the derived regional housing index can be just added to the RPI/CHAZ without the need to derive the cost of non-housing consumption by subtracting changes in the cost of housing from the National RPI, moreover this would give the same RPI as taking the CBSW and adjusting for regional variations.

The Regional RPI described in this section is essentially a new measure of the cost-of-living that partially updates the national RPI/CHAZ with regional housing index allowing therefore the original RPI to vary by regions. Housing cost is measured by price of gross rent⁶ from 1997 to 2008, derived from the household data from FRS carried out jointly by the ONS and the National Centre for Social Research⁷.

The categories of renters from which the cost of housing are derived are very similar to those considered by the ONS, such as households renting from councils; those renting from housing associations; and those renting privately, in both unfurnished and furnished accommodation.

Rental costs in the UK vary significantly across regions and over time. Since the 1990s, rental prices have been increasing consistently, particularly in London and the South East. Information based on the FRS shows that in 2008 the average nominal weekly household rent in London was about £163. This was about 57% higher than the average rent in Scotland (£70) in the

⁶ Consistently to the housing components for both the RPI/CBZW and the housing components used to derive the weights, gross rent includes council tax and rates, water and other charges, repair and maintenance and dwelling insurance.

⁷ Because data do not exist before 1997, it has not been possible to extend the analysis to earlier time periods.

same year and 47% higher than the rent in London in 1997 (£86). Housing costs not only vary considerably across regions in the UK but also they represent a large share of total expenditure; for example in 2008, housing costs made up 34% of total expenditures of people living in London and only made up 18% of people living in Northern Ireland (FES, 2008).

Wages are usually deflated using the national RPI which does not capture any regional variation in prices. The measurement of changes in real living standard requires nominal wages to be converted into real wages. To investigate the role of housing costs on wages of workers located in different parts of Britain, the measure used is the cost of housing, specifically gross rent faced by households in region r ($r=1 \dots 12$). In a similar vein to the methodology followed by Moretti (2010), the cost of housing used in this paper reflects the increase in the cost of housing experienced by individuals working in the same British region. Using gross rents has the advantage of being easy to measure and comparable to the ones used by the ONS in the construction of the RPI. To derive the regional RPI, the national RPI calculated by the ONS is partially updated by the cost-of-housing represented by gross rents, i.e. rent plus main charges.

To account for housing expenditure patterns in the 12 UK regions, the national RPI is re-weighted by appropriate regional plutocratic weights⁸ based on the same data (FES) and the same housing expenditure classifications used by the ONS to derive weights for the National RPI. Therefore weights are derived as the share of the total housing expenditure in total consumption expenditure in region r at the household level; the total housing expenditure includes rent, rates, water, council tax and other regular housing payments such as central heating repairs and maintenance; the total consumption expenditure includes the total housing expenditure, fuel, light and power; food expenditure; alcoholic drinks, tobacco; clothing and footwear, households goods, services, personal goods and services and motoring.

Deriving regional weights represents an improvement with respect to the methodology followed by Moretti (2010) that in fact does not derive local weights but uses the same national weights used to derive the CPI. In this paper, weights are derived by regions and for renters only; because the housing expenditure varies across regions so does the share of total expenditure. As underlined by the RPI technical manual (2007), “the RPI uses aggregate average expenditure to calculate weights implying that each index household contributes to the weights an amount proportional to its expenditure”. Therefore the weights affect the different items in different ways, hence the distribution of weights will differ.

⁸ The alternative approach is the democratic; democratic indices weight sample household equally and give straightforward means (arithmetic and geometric).

The cost of housing faced by a worker in a region is measured as the average of the weekly rent. The rationale for using rental costs is that rental costs are a better approximation of the user cost of housing; since houses are an asset, their prices reflect both the user cost as well as expectations of future appreciation (Moretti, 2010). The use of rents as proxy for housing costs is one of the approaches⁹ to the treatment of durable goods in a consumer price index or an RPI (Diewert, 2003) and corresponds to the rental equivalence approach. This approach to the treatment of durables is conceptually simple: it values the services yielded by the use of consumer durable goods for a period by the corresponding market value for the same durable for the same period of time (Diewert, 2003).

In a similar vein to the ONS methodology, the weighted sum of the cost of housing is normalised to 100 in 1997 and non-housing consumption normalised to 100 in 1997. The final Regional RPI (RRPI) can therefore be written as:

$$1) \quad RRPI_{rt} = (NRPI * (1 - w_{rt})) + (WH_{rt} * rent_{rt})$$

where r corresponds to the 12 regions in the UK; $(1 - W_{rt})$ captures the weight for non-housing consumption expenditure by region (r) and year (t), assuming that the cost of non-housing consumption is the same for all individuals in all regions. WH is the share of housing cost by region and year, and $rent$ is the index used as a proxy for cost of housing.

Another aspect that differs from the methodology used by Moretti (2010) is that while he uses the price for rent considering the monthly cost of renting a 2 or 3 bedroom apartment paid on average by graduate or high school graduate workers, in this chapter the rent is used irrespective of skill groups¹⁰, this is motivated by the idea to generate a house index that is more comparable to the one provided by the ONS that generates RPI irrespective of level of education or composition of the households. In order to compare consistently the price of rent across regions, average weekly rents derived are based on houses with 2 to 4 bedrooms.

Unlike Moretti, information for the price for rent is not based on individual data but is derived at household level. This is due to the lack of data at individual level; although potentially this could be derived from the EFS/FES/LCF that contains information on the level of education of households as well as the size of housing, due to the small sample the disaggregation by region would lead to imprecise estimates¹¹.

⁹See Diewert (2003) for a survey of alternative approaches.

¹⁰ This means that the measure used by Moretti will possibly add more variation to the final estimates.

¹¹ For example, in 2008 observations for graduate workers by regions ranged from 48 to 258.

How concentration of graduates and prices affect national estimates of wage dispersion.

The following table A provides an illustrative example about how the notional concentration of graduates in one area is likely to affect estimates of wage dispersion. The example uses hypothetical weights as a share of graduate and non graduate workers as well as hypothetical nominal weekly wages (in Pounds) and Regional RPI.

Supposing there are two regions, Region 1 and Region 2, characterised by workers of only 2 types of education (graduate and high school) unevenly distributed between the two regions; let us assume that Region 1 has a higher concentration of graduates (0.4) and a lower concentration of high school graduates (0.1) and is more expensive than Region 2 with a regional RPI equal to 1.4; by contrast Region 2 has a higher share of graduates (0.6) but also a higher share of high school graduates (0.9) and is relatively less expensive than Region 1 with a regional RPI equal to 0.95. Let us also assume that both graduate and high school graduate workers in Region 1 earn more than their counterparts in Region 2. The calculated weighted average graduate–high school wage gap decreases by 14% (0.20) when using the Regional RPI with respect to the nominal wage. In example 2, the Regional RPIs used (column 6) are the RPIs for London in 1997 as a proxy for the RPI in Region 1 and the average RPI in all remaining UK regions in the same year as a proxy for the RPI for Region 2. In this case, the nominal weighted average graduate-high school wage gap is only 5% (0.08) higher than the regional one; example 3 keeps the hypothetical share of workers by education and the RPI in 1997 but imposes the same nominal weekly wages in the 2 regions allowing only to differ by education. In this case the difference between the graduate-high school wage gap using the nominal wage and that one obtained deflating the nominal wage by the regional RPI decreases even more, being only 4% (0.05) higher.

This simple example illustrates how the effects of the regional deflation on the average national wage may differ in magnitude depending on the combination of other factors. Specifically, the final effects of the regional deflation are likely to be determined by the combination of at least three elements: the concentration and distribution of graduates/non graduates across regions; the difference in prices and therefore the difference between the cost-of-living of different regions and the regional difference in the wages of graduate/non graduate workers. The less difference that exists between those three elements across regions, the smaller will be the effect of the regional deflation.

Table A: Illustrative example

	Share Graduate	Share High school	Wage Graduate	Wage High school	Regional RPI
Example 1.					
Region 1	0.4	0.1	200	120	1.4
Region 2	0.6	0.9	150	100	0.95
Average Based on Nominal wage			170	102	
Graduate- High school gap (as ratio)				1.67	
Average based on regional deflated wage			151.9	103.3	
Graduate- High school gap (as ratio)				1.47	
Change in the National-Regional wage gap				0.2	
Example 2.					
Region 1	0.4	0.1	200	120	1.13
Region 2	0.6	0.9	150	100	0.98
Average Based on Nominal wage			170	102	
Graduate- High school gap (as ratio)				1.67	
Average based on regional deflated wage			162.6	102.5	
Graduate- High school gap (as ratio)				1.59	
Change in the National-Regional wage gap				0.08	
Example 3.					
Region 1	0.4	0.1	150	120	1.13
Region 2	0.6	0.9	150	120	0.98
Average Based on Nominal wage			150	120	
Graduate- High school gap (as ratio)				1.25	
Average based on regional deflated wage			144.93	120.82	
Graduate- High school gap (as ratio)				1.20	
Change in the National-Regional wage gap				0.05	

Table B shows how those three elements interact and determine what we observe in the UK. The table uses data from LFS for 1997, considering separately London, the South East and the remaining British regions as a single region and provides an idea about how the regional deflator affects the graduate wage gap for the two regions separately and what is the final effect for the whole of the UK. The table reports the share for each region with respect to the total population

(column 1); the regional deflator for London, South East and the remaining regions (column 2), the share of graduates, high school and less than high school workers in each region (columns 3, 4 and 5 respectively), their hourly wage in both real and nominal terms (in brackets) (column 6, 7 and 8) as well as the graduate-high school and graduate less than high school gap reported both as difference (column 9 and 10) and ratio (column 11 and 12). The first difference worth noting is that London had a higher Regional RPI in 1997 equal to 113.55, while the RPI for the South East was lower (105.59) but still higher than the average Regional RPI of the remaining regions (97.22). This is shown in column 2. These differences are due to the regional variation in the cost of housing. London and the South East also present the higher share of graduates (22% and 16% respectively); moreover the share of less than high school graduates in London comprises only about 8% of the working population.

The last column reports the graduate wage gap separately in London, the South East and the remaining regions in the UK. The example below shows that despite the graduate high school wage gap in London being lower than that in both the South East and the rest of the UK, the overall effects of the latter take over giving a final average wage gap that is higher than that of London and the South East considered separately. As shown in column 8, deflating wages by the Regional RPI decreases the absolute graduate high school wage gap by about 11% in London and 4.8% in the South East, while it increases by about 2.4% in the rest of the UK. The overall effect at a national level of the Regional RPI will be a decrease of about 4.2% of the graduate high school wage gap compared to the nominal gap. A similar pattern can be observed when looking at the graduate less than high school wage gap reported in column 9; the use of the regional RPI decreases this by 11% in London, and by about 5.6% in the South East, while it increases by about 2.4% in the rest of the UK resulting in an average national decrease of about 4.2%. Columns 11 and 12 also report respectively the graduate –high school wage gap and graduate less than high school wage gap; however the use of the regional deflator does not make any change to the ratio compared to the one derived using the national deflator.

The illustrative example together with the table reporting evidence from 1997 documents that the observed regional dispersion of graduate and non-graduate shares, together with the relatively small differences in price levels (captured by the Regional RPI) as well as the relative small difference in the hourly wage earned by graduates in London and the in the Rest of the UK while make more difference to estimates of real wage levels for subpopulations concentrated in certain (more expensive) areas, will result in an overall national effect on wage gaps that is almost zero.

Table B: Illustrative example.

	Pop. share	Reg. RPI	G share	HS share	LHS Share	G Hourly wage	HS Hourly wage	LHS Hourly Wage	G-HS Wage gap	G-LHS Wage Gap	GHS Wage Ratio	GLHS Wage Ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
London	11.7	113.55	22	14.3	7.6	11.22 (12.69)	8.18 (9.27)	6.6 (7.50)	3.04 (3.42)	4.62 (5.19)	1.37 (1.37)	1.70 (1.69)
South East	14.7	105.59	16	17.2	13.2	12.20 (12.83)	8.46 (8.90)	6.67 (7.00)	3.74 (3.93)	5.53 (5.83)	1.44 (1.44)	1.83 (1.83)
Rest of the UK	73.6	97.22	62	68.5	79.2	11.43 (11.16)	7.64 (7.47)	6.37 (6.22)	3.78 (3.69)	5.06 (4.94)	1.50 (1.49)	(1.79) (1.79)
National (average) wage gap						11.55 (11.76)	7.87 (7.97)	6.43 (6.42)	3.64 (3.80)	5.12 (5.34)	1.47 (1.48)	1.80 (1.83)

Notes: Based on the LFS and EFS/FES/LCF for 1997, nominal wage in brackets. G = graduate; HS= high school graduate; LHS = less than high school graduates.

3. Identification and estimation

Because the aim of this chapter is to investigate if and how much regional variations in the cost-of-living can account for changes in both the graduate high school wage gap and the 90-10 wage gap between 1997 and 2008, the econometric methodology estimates both the conditional nominal and real wage difference between workers with a graduate and high school degree as well as the difference between workers with a college or more degree and a less than a high school degree. The baseline estimates are based on a regression of the log hourly wage, nominal and real separately, on a dummy variable indicator for graduates interacted with a dummy for each year.

Controls include race, gender, year dummies, and a cubic for potential experience. Other regressions are run adding regional fixed effects.

As explained by Wooldridge (2006), when wages appear in logarithmic form and dummy variables are used for all time periods, the use of aggregate price deflators will only affect the intercepts but will make no difference for the slope estimates. In fact when the log wage is used as a dependent variable, provided that a year dummy is included in the regression, using real or nominal wages will only affect the coefficient of the year dummy¹².

¹² Following Wooldridge (2006) suppose to deflate wages of 2008 at 1997 denoting the deflator factor for 2008 wages P_{2008} (1.31 using RPI =100 in 1997) then the log of real wage for each individual in the 2008 sample can be written as: $\log(\text{wage}_i/P_{2008}) = \log(\text{wage}_i) - \log(P_{2008})$.

To classify education, the variable “age left full time education” is used. Following Manacorda, Manning and Wadsworth (2007), the main estimates are based on three education groups defining anyone who left full time education at the age of 16 or less as “less than high school”; anyone who left full time education between the ages of 17 and 20 as “high school graduate”; and anyone who left education at the age of 21 or later as “college graduate”. In order to keep the analysis as consistent as possible with Moretti (2010), the sample for the baseline estimates includes all UK born workers, aged 25-60 working both full time and part time. However, additional specifications extend the analysis to all of the working age population.

The baseline regression specification can be written as follows:

$$2) \text{ Log } w_{iT} = \delta_0 + X_{iT} + C * Y_t + Y_t + \varepsilon_{iT}$$

Where w_{iT} is the nominal hourly wage for individuals i in year T ; X_{iT} includes a set of controls such as cubic in potential experience, gender, race; C is an indicator for college interacted with each year; Y_t is a year dummies and ε is an error term.

The baseline specification for the real wage changes in the dependent variable that now can be written as the nominal wage divided by the RPI_r where $r = 1, \dots, 12$, therefore the specification to estimate the conditional real wage difference between graduate workers and high school (less than high school) can be written as:

$$3) \text{ Log } (w_{iT}/RPI_{rT}) = \delta_0 + X_i + C * Y_t + Y_t + \varepsilon_r$$

Additional specifications also add region fixed effects to the baseline regressions.

4. Data

The empirical analyses are based on the combination of three datasets: The Expenditure and Food Survey (EFS), (renamed Living Costs and Food Survey (LCF) in 2008 and formerly the Family Expenditure Survey (FES)) provide the information used to construct regional housing weights. The Labour Force Survey (LFS) is the primary source for individual earnings and education data and the Family Resources Survey (FRS) from 1997 to 2008 provides the information

Because wages differ across people but P2008 does not, the log(2008) will be absorbed into the intercept of 1985. However this conclusion will change if the RPI differs for people living in different parts of the country.

on housing price rents across regions. The lack of information for rental prices for all regions before 1997 does not allow the analysis to be extended to an earlier time period.¹³

The EFS/FES/LCF is used to derive weights and is the same data source that the ONS uses to derive the RPI. The EFS/FES/LCF is a continuous cross-sectional survey that has been carried out by the Office for National Statistics (ONS) since 1957 which monitors the spending patterns of around 6,000 to 6,500 household across the country each year.

The FES ran from 1957 to March 2001. From April 2001 onwards, the data continues to be collected in the EFS, formed by combining the FES with the National Food Survey (NFS). The LCF is a continuous survey conducted by the ONS, which moved from a financial year to a calendar year basis from 2008. In the EFS/FES/LCF, households are sampled from randomly selected postcode sectors stratified according to region across the UK, car ownership and socio-economic status.

There are two major components to the survey. A two-week paper-based diary that records all expenditures and an interview that collects information on household demographics, income and some retrospective information on regular purchases (such as rent, mortgage payments and utility bills) and irregular, expensive purchases (such as durables and holidays). Expenditures are calculated and recorded as household-level weekly averages in a number of relatively disaggregate categories- for food there are around 100 such categories. Data is collected throughout the year to cover seasonal variations in expenditures. In addition to expenditure and income data, the EFS/FES/LCF collects information on socio-economic characteristics of the households, e.g. composition, size, social class, occupation and age of the head of household.

However the EFS/FES/LCF has a number of drawbacks. One is that it does not cover all households such as people living in retirement homes, military barracks or student halls of residence or residents in temporary homes. Another problem is mainly due to the response rate. Around one-third of those initially approached do not respond to the survey.

Established in 1973, the LFS is the largest survey of households living at private addresses and in NHS accommodation in the UK, conducted by the Office for National Statistics (ONS). Since 1992, the LFS has been a rotating quarterly panel. Information is recorded in four quarters; each quarter's LFS sample of 53,000 UK households is made of five "waves" each of approximately 11,000 private households. Each wave is interviewed in five successive quarters, earnings information is only recorded in waves 1 and 5. A single stage sample of addresses with a random start and constant interval is drawn from the Postcode Address File (PAF) sorted by

¹³ Similarly when this paper has been started data to derive weights (LCF) for 2009 were not released yet.

postcode. The LFS also contains information at regional level, where region is determined according to usual residence. The LFS identifies 20 regions¹⁴. These 20 regions are unified to be consistent with the 12 more limited regions identified in the EFS/FES/LCF. These comprise the North East, North West, Yorkshire, East Midlands, West Midlands, East, South East, South West, London, Wales, Scotland, Northern Ireland. The LFS contains detailed information on individual characteristics, age, marital status, migration status, job characteristics, wages and hours worked. It also contains information on the housing tenure of the individuals; giving on average 11000 observations of graduates reporting a non zero wage in each year and between 38000 and 58000 non-graduate workers with a positive wage.

The data used for the price of rents are based on the FRS from 1997 to 2008; the FRS is a continuous survey with an annual target sample size of 24,000 private households. Fieldwork is carried out jointly by the Office for National Statistics and the National Centre for Social Research. The survey was launched in October 1992 to meet the information requirements of the Department for Work and Pensions (DWP) analysts. Households interviewed in the survey are asked a wide range of questions about their circumstances, including receipt of Social Security benefits, housing costs, assets and savings. Before 2002 the survey for Northern Ireland was carried out for the Department for Social Development (DSD) by the central survey unit of the Northern Ireland Statistics and Research Agency (NISRA) of the department of finance and personnel.

The price for rent used to construct the regional RPI is the total amount of rent eligible for Housing Benefit paid by a household before the deduction of any housing benefits. In particular, the tenants used to derive information are only those renting privately, therefore tenants who are in rent free accommodation are excluded from the sample. Although these are rents for low income households and may possibly under-estimate the costs for graduates, they are very similar to the rents used by the ONS in the construction of the RPI (CBZW).

There are at least three other potential data sources that could be used as proxy for housing costs to construct a Regional RPI. The first one is the private rent available from the Department for Communities and Local Government based on the Survey of English Housing (SEH). The SEH presents a few limitations: first of all figures available are mean rents over two financial years, for example 1997 figures are a mean of the 1996-1997 and the 1997-1998 rents. Secondly, from 2008, information at regional level is no longer available; finally the SEH only covers English regions,

¹⁴ The 20 regions indicated in the LFS are the following: Tyne and Wear; Rest of North East; Greater Manchester; Merseyside; Rest of North West; South Yorkshire; West Yorkshire; Rest of Yorkshire & Humberside; East Midlands; West Midlands Metropolitan County ; Rest of West Midlands; East of England; Inner London; Outer London; South East; South West; Wales Strathclyde; Rest of Scotland, Northern Ireland.

therefore excluding Wales, Scotland and Northern Ireland for which there is no other comparable information on private rents.

The second possible data source is represented by the regional housing price available from the Nationwide Building Society that covers quarterly house price changes from 1973 to 2008. The construction of a regional RPI requires data to be comparable with housing costs used by the ONS, however the housing prices available from the Nationwide are neither fully comparable to the RPI using mortgage interest or the RPI using rents. Moreover, the Nationwide definition of regions differs in significant ways from the ONS definition.¹⁵

The third possible data source available for the UK is based on the Registered Social Landlords (RSLs) and is equivalent to the local authority rents data. RSL rents are derived from the Regulatory and Statistical Return that the Tenant Services Authority (TSA) sends out once a year to all RSLs. One concern related to the use of the RSL rents is that they tend to be for low income households and are therefore likely to under-estimate housing costs, particularly in London.

Rents from the FRS are available for all 12 regions and they reflect the differences across UK regions and the increase over time; the FRS remains the more comparable housing cost with those used by the ONS for the construction of the national RPI.

The full samples used for the empirical analysis are men and women aged respectively 16-64 and 16-59, though the baseline estimations replicating Moretti's focus on individuals between 25 and 60, so the analysis is limited to workers who are employees, both full time and part time; considering only their main job, and who report a positive wage. To limit the effect of outliers, following the existing literature in the UK (Manacorda, Manning and Wadsworth 2007) only observations with an hourly wage between one and a hundred pounds at 1997 levels are used.

5. Results

5.1 Descriptive statistics

This section begins with some descriptive evidence on the distribution of workers by level of education (graduate, high school and less than high school) across UK regions and over time followed by a discussion on the differences in levels and measures of dispersion of hourly and log wage. The sample is based on the working age population (16-59 for women and 16-64 for men) working both full-time and part-time, employees and considering the main job only.

¹⁵ See Holly, Pesaran and Yamagata (2010) for details on difference in regional definitions.

Figure 1 together with Table 1 describe the distribution across UK regions of workers by level of education in 2008. Graduate workers are not evenly distributed across the British regions and are overrepresented in the South-East and in particular in London. This is shown by figure 1. In 2008 about 38% of the National graduate work force was concentrated in two regions: London (21.1%) and the South East (16.2%). These two regions also experience the higher concentration of the total working age population, respectively 11.5% and 15.0%. As documented in Table 1, Northern Ireland and the North East hosted the lowest shares of graduate workers, 1.8% and 3.0% respectively, while 1.7% and 4.4% of the total working age population were concentrated in the two regions respectively. The high school graduates distribution in London and the South East is in proportion with its population share, with the highest percentage of high school workers being concentrated in the South East (16.5%) and London (11.4%). Only 6.2% of the “less than high school” workers are concentrated in London.

Over the sample period the share of graduates in the UK population rose from 15% in 1997 to 23% in 2008, with an average increase over time of 56%. Figure 2 looks at the changing regional concentration of graduates over the sample period, giving the graduate share in each region for 1997 and 2008. It shows that the share of graduate workers has been increasing over time in all 12 regions, particularly in London where the share of graduates increased by about 53% between 1997 and 2008. This, coupled with the fact that, as documented in Table 3, in the same time period the average cost of housing in the UK has been increasing by 79% with the increase being higher in London (89%) implies that graduate workers are more likely to face a higher cost-of-living than in 1997 and to experience relatively greater rises in housing related costs than other education groups.

Table 2 reports the percentage distribution of workers by level of education in the 12 UK regions in 1997 and 2008. In 2008 nearly 43% of the London work force had a graduate level of education compared to 29% in 1997; the South East is the second region with the larger percentage of graduates specifically 17% in 1997 and about 27% in 2008. On the other hand the North East reports the lowest share of graduates both in 1997 (10%) and 2008 (17%), while in 1997 about 71% of the North East working population had a less than high school level of education; almost double the share in London in 1997 (37%). Between 1997 and 2008 there has been a relative convergence in the level of graduate workers across the UK regions with the higher percentage increase experienced by Wales (76%), Yorkshire (64%) and North East (61%) and the lowest (39%) by the East Midlands. Despite the change in this relative concentration of graduate workers, the absolute concentration of them in specific regions (London and the South East) remains the major focus of this study because the absolute shares matter for housing calculation.

The picture painted about the distribution of workers by education across the UK regions implies that the greater part of graduate workers is concentrated in the more expensive regions: London and the South East.

Based on the FRS, Table 3 reports weekly nominal rental prices across British regions in 1997 and 2008, the increase from 1997 to 2008 (column 3) and the percentage increase relative to the national mean (column 4). The table documents that the populations of London and the South East experience higher rents compared to both the UK average and the remaining regions and this is persistent over time. In 2008 the average weekly rent was about £163 in London and £123 in the South East while the UK average rent equalled £92. Northern Ireland and Scotland remain the “cheapest” regions in terms of housing costs. Looking at the UK as a whole between 1997 and 2008, the nominal price for rent increased by about £41 with an average change of 79%; the increase in price for rents in London was 89% (£77) higher than the national one while in the North East the increase was about 28%, lower than that of the UK. These statistics coupled with those reported by Figure 1 and Table 1 suggest that graduate workers in London and the South East are more likely to have experienced both a higher cost-of-living and a more rapid increase in costs due to their higher housing expenditure over the sample period implying that the relative increase in their real wages might be smaller than those of counterparts living in less expensive regions.

To better understand if and by how much the national and regional RPI diverge, Figure 3 plots the trends of the two series (National RPI and Regional RPI) by region in the UK from 1997 to 2008. The figure clearly shows the remarkable differences between the two RPIs for London; the higher level of the regional RPI in London clearly illustrates that the actual cost-of-living in London is higher than what the national RPI demonstrates, other things equal (i.e. accounting for housing – but not other costs which may be higher (or lower) in London. These results are consistent with Borroah et al. (1996) who demonstrate that the relative expensiveness of London and the South East, over 1979-90, increased when housing cost was included. The Regional RPI in London is not only higher than the national one but is also the highest across the regions. The difference between the two trends can be interpreted as due to the (higher) cost of housing in London coupled with the higher share of housing expenditure over the total expenditure. The second region for which the regional RPI highly diverges from the national one is the South East; this is not surprising given that this is one of the regions with the highest price for housing.

The opposite is true for the remaining regions, particularly for Scotland and the North East that appear to face a lower cost-of-living than the one represented by the national RPI. Less difference between the two RPIs is found in the East and South East, where the gap is less evident.

Table 4 reports real hourly wages for all workers, as well as men and women separately, in 1997, 2000, 2003, 2006 and 2008 when using different measures of inflation for the nominal wage deflator. Column one reports the nominal hourly wage; column two reports the hourly wage deflated by the National RPI provided by the ONS; column three reports the hourly wage deflated by the Regional RPI. The table documents that the real hourly wage deflated by the Regional RPI is lower than that deflated by the national one until 2000 but higher after 2000. As explained later this effect is likely to be due to the fact that until 2000 deflating wage by the Regional RPI generates average lower wage in all regions, however after 2000 this is only true for London and the South East with the remaining regions experiencing higher Real (Regional) wage that overcome the lower wage in London and the South East. In fact from 1997 to 2000 deflating hourly wages by the Regional RPI the real hourly wage for all workers decreases from 6 (8%) to 12 (13%) pence an hour with respect to the hourly wage deflated by the National RPI while after 2000 it increases from 10 (11%) to 27 (30%) pence an hour with respect to the hourly wage deflated by the National RPI. However, it is important to stress that the differences in the hourly wage deflated using the National RPI and the one deflated using the Regional one are not that great. On average men experience a higher loss in wages due to the regional cost of housing though the difference is negligible.

Because London and the South East are the most expensive British regions, it is reasonable to consider these regions separately in contrast to the other UK regions. Table 5 reports the hourly wage deflated by using the National RPI and the Regional RPI from 1997 to 2008 for the whole of the UK, London, the South East and all the other remaining regions together. There is now a clear trend that better explains the decrease of regional wage up to 2000 and the decrease afterwards; in fact from 1997 to 2008 in London and South East the hourly wage deflated by the Regional RPI is always lower than that deflated by the National RPI but it is always higher in the remaining regions. The hourly wages deflated by the Regional RPI are particularly lower in London; on average the decrease in real hourly wage due to the cost of housing ranges between 1.10 (11%) to 2.18 (18%) pence in London and 0.22 (2%) to 0.57 (6%) pence in the South East¹⁶; by contrast in the remaining regions the hourly wage deflated by the Regional RPI is on average higher 15 (2.2%) to 81 (9.4%) pence. These figures imply that the National RPI is likely to underestimate the cost-of-living in London and the South East while overestimating it in the remaining regions.

To consider how different measures of inflation can also affect the measure of wage dispersion, Table 6 reports some common measures of wage inequality based on the log hourly

¹⁶ Leunig and Overman (2008) provide the theoretical justification to this explaining that in terms of living standards there exists an optimal city size: in practice if cities are larger or smaller than the optimum then productivity, wages and employment will be lower than they could be.

wage for all workers for the years 1997, 2000, 2003, 2006 and 2008. Column 1 reports wage dispersion for wages deflated by the National RPI; column 2 reports similar measures for wages deflated by the Regional RPI, column 3 reports the difference between column 2 and column 1. The measures used are standard deviation, variance, the 90-50 gap, 90-10 gap and 50-10 gap.

There are a few aspects that are worth noting. When using a regional deflator, there are hardly any changes with respect to the national one; in fact using the Regional RPI the average changes in wage inequality from 1997 to 2008 decreased from 0.006 to 0.016 log points with respect to the national deflator, using as a measure of dispersion either the standard deviation or the variance. The differences are slightly more notable when looking at the 90-50, 90-10 and 50-10 wage gap, though still not striking. Using the difference between the 90th and 50th percentile as a measure of dispersion, when deflating wages by the regional RPI, the difference decreases from 0.011 to 0.019 log points with respect to the national deflator; the difference for the 50-10 gap ranges between 0.005 and 0.021 log points and while this is notable, slightly more difference can be seen in the 90-10 gap. Using the regional RPI to deflate wages, the 90-10 wage gap decreases from 0.019 to 0.033 log points with respect to the national deflator. However the differences are not substantial. For example, in 2006 the 90-10 wage gap for all workers was equal to 1.349 when using the national deflator and decreases to 1.316 when using the regional one. Panel B of table 6 also reports the changes in wage inequality from 1997 to 2008; the panel documents that from 1997 to 2008 LFS estimates of wage inequality in the UK decreased. This is true for all measures used except for the 90-50 for which there has been almost no change. For example, over the time period analysed, the standard deviation decreases by 0.021 log points, while the highest decrease can be observed for the 50-10 wage gap (-0.051). Using the regional deflator does not make much difference to those measures both in terms of sign and in terms of magnitude.

Because the 90-10 gap can be used as a proxy for the graduates-less than high school gap, given that the number of graduates is concentrated in areas (London and the South East) that are more expensive and therefore as shown in Table 5, those are the areas whose costs are likely to be underestimated by the current RPI provided by the ONS. This raises the question on whether the construction and therefore the use of a regional RPI would be more appropriate and representative of the real cost-of-living. Table 7 presents similar measures of dispersion for the 12 regions in the UK in 2008. While, by construction, these within-region measures reported are the same when using any measure of inflation (National RPI or Regional RPI), it documents the variation of wage inequality across the UK regions. The first relevant aspect is that the more expensive regions (London and the South East) are also the regions where wage inequality is higher than the national average. For example in 2008, considering the standard deviation as a measure of wage dispersion,

London experienced approximately an 8% higher wage dispersion (based on the standard deviation) than the national average; the dispersion increases (14%) when looking at the 90-10 gap in London compared to the UK average suggesting also that graduate workers in the London area experience higher wage dispersion than their counterparts in the rest of the UK, higher for example than the North East and Northern Ireland. This highlights the importance of regional specific studies on wage inequalities.

Table 8 compares regional changes in the level of the real hourly wage from 1997 to 2008 for all workers, when deflating wages by respectively the National RPI and the Regional RPI. The table demonstrates two main facts: the first is the variation in changes in real hourly wages within UK regions. Based on the national RPI, column 1 documents that when deflating the real hourly wage by the national RPI London experienced the highest increase (£2.43) corresponding to 53% higher than the national one (£1.59) while the East Midlands experienced the least growth (£1.15) in real hourly wages corresponding to 28% less than the average increase in real hourly wages. Column 2 shows how those changes vary when deflating the real hourly wage by the appropriate Regional RPI that accounts for different levels of rent. Although the within variation by region in changes remains, there are two main new trends: in regions where the cost of housing is typically higher (i.e. London and the South East) the estimated changes in regional real hourly wages over time become smaller. In London the real hourly wage increase is £1.37 compared to £2.43 based on the National RPI and is now 29% lower than the average UK increase (£1.93); similarly in the South East the real hourly wage change (£1.61) is now 17% lower than the UK average. This implies that because graduates are more concentrated in those two regions, and due to the higher rises in living expenses, their real hourly wages have been increasing by less than the UK national average rate.

The opposite occurs for the rest of the regions. When accounting for regional specific rents the change in real hourly wages between 1997 and 2008 is higher. Scotland is now facing the highest gain (61 pence) with respect to the UK average, corresponding to a real hourly change between 1997 and 2008 32% higher than the UK.

These findings confirm that the national measure of inflation that does not account for regional variations in the cost-of-living can affect estimates of real wage growth in local areas. The previous description about how level of wage changes depending on the measure of inflation used (i.e. regional or national) suggests that the actual National RPI provided by the ONS may not reflect the actual level of prices and therefore the actual cost-of-living faced by differently skilled workers of different British regions.

To document this, Table 9 shows the average weekly rent for 1997 and 2008 by level of education for the UK, London and the South East. In both years considered, graduate workers paid a higher weekly rent than the high school graduate and less than high school graduates. For example, in 1997, graduates' weekly rent equalled £98.31, while high school and less than high school workers were paying respectively £65.77 and £46.30, when looking at graduates in London the rent rises to £112.40. Similarly in 2008, graduates in the UK were paying an average weekly rent of £164.39, higher than both the high school (£105.72) and less than high school (£84.58). When focusing the analysis on graduates located in London, these differences are amplified with graduate weekly rent now being around £244.2, almost double that of the less than high school graduates in the same area. As explained in the data section, those rents are the total amount of rent eligible for Housing Benefit paid by a household before the deduction of any housing benefits. The tenants used to derive information from are only those renting privately, those renting from Landlord Associations and from Councils, however tenants who are in rent free accommodation are excluded from the sample. The last three columns of table 9 report the percentage increase for graduate, high school and less than high school respectively in the UK, London and the South East. The table documents that graduate workers in London experienced the highest increase in weekly rent (117%) from 1997 to 2008, compared to the average percentage increase for all graduates in the UK (67%).

The relationship between increasing share of graduates and increasing price for rent is shown in figure 4 reporting the average rent by region and year from 1997 to 2008 in the horizontal axis and the share of graduates per region and year from 1997 to 2008 on the vertical one. The positive relationship indicates that regions that have experienced the largest increase in the share of graduates are the regions where the average cost of housing is higher and increased the most.

To observe how graduate workers are more likely to face a higher and rising cost-of-living affecting the real wage, Table 10 gives the absolute real hourly wage differences for college and high school graduate workers, for male and female when deflating the hourly wage by the National or the Regional RPI. The table shows that the observed differences are very small and not significantly different. Panel A reports changes for graduate men and women; on average despite the persistent gender gap, graduate men and women face a lower hourly wage when accounting for regional cost of housing; this is true except that women in 2008 gain 7 pence when deflating wages by the Regional deflator Panel B reports similar results for high school graduates; in this case the difference between hourly wage deflated by national RPI and regional RPI is positive in relative terms only in 1997 but there is a gain when using the regional RPI. Table 11 gives more details of the real hourly wage for all workers by 12 UK regions in 2008 using the National and Regional RPI documenting that when looking separately at regions the hourly wage for graduate workers in

London and the South East decreases when deflating wages by the Regional RPI while it increases in the remaining regions.

The trends and levels of the hourly wage for graduates and high school when using the National or Regional RPI are better understood from figures 5a and 5b that display the time series of the real hourly wage respectively for graduate and high school for all workers by regions from 1997 to 2008. For each region, the real hourly wage deflated by the national RPI is graphed alongside the real hourly wage deflated by the Regional RPI. Although graduate workers in London earn on average more than other graduates in the rest of the UK they also clearly face a higher cost-of-living than the one reported by the National RPI earnings, therefore a lower wage in real terms. This difference is persistent and increasing over time due to the increasing cost of housing affecting real wages in London more than it does in other regions. Similarly, for workers in the South East, the real wage is lower than the national real wage. Graduate workers in Scotland have the advantage of a lower cost of housing and so experience a higher real wage than the one actually determined by the ONS. While there is not much difference for the East Midlands, West Midlands and South West, in the remaining regions graduate workers earn more. The difference between real wages deflated by the National and Regional RPI is qualitatively similar when looking at the high school workers (figure 5b), though the gap is lower.

Figures 6a and 6b compare the trends in real hourly wages deflated by National and Regional RPI for male and female graduates and high school graduate workers in London and in the UK as a whole. Figure 6a plots the trend for real hourly wages for male graduates and high school working in London and in the UK as a whole; while there is not much difference when looking at the average real hourly wage for the UK the gap persists within and versus London. This is more remarkable for graduates; in fact the figures show that graduate workers “lose” more in absolute terms compared to high school due to the housing costs in London though the main trend remains. Figure 6b reports similar results for females. The figures highlight the persistent gender wage gap within both graduate and high school workers, but also show that the hourly wage has been increasing for graduates more than for high school graduates.

Figure 7 plots the wage difference between graduates and high school and graduates and less than high school for all workers in the UK using the National RPI and the Regional RPI. The dashed navy line plots the wage gap when using the regional RPI; the figures show that both graduate-high school and graduate-less than high school wage gaps decrease over time and are lower when using the regional RPI that accounts for the housing-cost-of-living compared to the gap based on the National RPI. Table 12 reports the difference between the wage gap based on the National RPI

and the one based on the regional one for graduate-high school and graduate-less than high school workers. On average, the difference is higher for the graduate-high school (column 6).

5.2 Estimation results

This section presents the estimates of to what extent accounting for these spatial issues affects estimates of the changing returns to education, beginning with an exploration of how much of the changes in nominal wage differences between graduates and high school observed in table 11 are due to regional differences in the cost of housing in the UK. Table 13 replicates table 4 of Moretti (2010). Model 1 estimates the conditional nominal wage difference between college graduate workers and high school. All estimates are from a pooled sample containing observations from 1997 to 2008 based on a regression of the log nominal hourly wage on an indicator for college interacted with a indicator for each year, year dummies, a cubic in potential experience, and dummies for gender and race. In order to compare the estimates with those of Moretti, the sample includes workers aged 25-60, who are UK natives working both part time and full time. The coefficients given in the table are the college-year dummy interaction terms from 1997 and 2008 and represent the conditional wage difference for a given year.

In 1997, the nominal gap is 1.4 % higher than the real gap and 1.8% higher in 2008. Column 3 reports the difference in the estimates between 1997 and 2008, and indicates that the conditional nominal wage difference between workers with a high school degree and workers with college or more has decreased by 0.017 log points over the period. The conditional difference between the wage of graduates and high school graduates decreases by only an additional 0.05% (from 0.017 to 0.018) when using the regional deflator¹⁷. However this fall is not statistically significant. This stands in contrast with the main estimates of Moretti (table 4, 2010) who reports that in the US between 1980 and 2000 the conditional nominal wage difference between workers with a college degree and a high school degree increased significantly. Moretti demonstrates that the conditional difference between the wage of college graduates and high school graduates is .60 in nominal terms and only .53 in real terms when the Local CPI is used as a deflator.

¹⁷ Because the within-group variance explains the most part of the increase in wage inequality (see Rienzo 2010 for an extensive review) it is also relevant to analyse the inequality in the within-group. Estimates from model 2 based on the Nominal wage show that the within-group variance increased from 0.234 in 1997 to 0.249 in 2008 with a difference of 0.015; when looking at the within-group derived by the model 2 that accounts for the Regional cost-of-living the level of the within-group variance increased from 0.228 in 1997 to 0.244 in 2008 ; though the level is slightly lower there is not much difference in the increase, and unlike the other measures of inequality presented earlier and in the estimates, the within-group variance has been increasing over time.

The difference between the estimates for the UK and those reported by Moretti for the US can be explained by any differential regional changes that jointly affect the final national estimates.

The reasons can be related to the combination of three elements, as explained earlier in the illustrative example A: the different concentration of graduate/non graduate workers across less and more expensive regions; the difference in the regional RPI; the difference in the hourly wage of workers in more expensive regions with respect to less expensive regions.

While Moretti (2010) reports that in 2000 in some of the US metropolitan areas the largest share of workers with a graduate degree among their residents was 58%, this was about 5 times the fraction of the college graduates in cities with the lowest share. Based on the LFS in the same year, 2000, London is the only area with the highest share of graduate workers (about 32%) which is on average only twice the share of graduates in any other region. The share of graduates in London has increased over time (reaching 43% in 2008) but so did the share of graduates in all regions. Similarly, the relative shares of UK graduates concentrated in London remained relatively constant over the sample period, at around 20%.

Another aspect that could help to explain the difference between the UK and US estimates can be related to the cost-of-living. Moretti documents that between 1980 and 2000 the housing costs, measured by the monthly rent, for graduate and high school workers increased by 147% and 127% respectively. Data based on the FRS shows that between 1997 and 2008 in the UK the housing costs, measured by weekly rent, for graduates and high school graduates increased on average respectively by 67% and 61% for the whole UK and by 117% and 87% for London, while it increased more (83%) for less than high school workers in the UK and 84% for those in London. This will result in a smaller difference between the National and the Regional RPI and therefore will reflect in a smaller effects of the regional deflation.

Additionally, data for the price of housing based on Nationwide shows that the increase in price of housing between 1980 and 2000 has been higher (245%) than the similar increase between 1997 and 2008 (185%), implying that the time period analysed by Moretti because characterised by a higher increase in price of housing, will also reflect in a bigger difference between the national and the regional measures of inflation. Another element that may have a role for the estimates can be related to the relatively small difference between the real hourly wage of workers, particularly graduates, in London and the rest of the country. As documented by Table 12, based on the National RPI between 1997 and 2008, the wage gap of graduates to high school workers decreased by 0.21 pence an hour (5%); this contrasts with the pattern depicted for the US between 1980 and 2000 by Moretti (2010). Moreover, using the regional deflator in 2008 the wage gap is only 9 pence (2.4%) higher than the 1997 one. Despite the fact that the regional RPI deflator decreases (London

and the South East) or increases (rest of the UK) the real wage level in 12 years (1997 -2008) there is very little difference between the two sets of estimates of the change in wage inequality (based on nominal and regional deflator).

It is likely that the picture for the US as painted by Moretti may rely on the existence of more variation for the rental prices and a higher proportion of graduates in more expensive areas that will affect the final estimates. The crucial condition that allows us to estimate a real hourly wage model like model 2 is that the price deflator varies each year.

For the UK, the more expensive areas made up only the 25% of the sample used, while for the remaining 75% the use of the Regional RPI translates into an increase rather than a decrease in real wage. This might possibly explain the fact that deflating wage by regional RPI makes little difference to the estimates of the relative wage gaps of graduates versus high school workers.

Another reason may be related to the fact that the cost of housing used in the construction of the index here may not fully capture differences in regional prices in the UK. London, together with the South East, is the region experiencing the higher share of weekly expenditure and higher costs not only for housing but also for transport and recreation (Family Spending 2009, ONS). Baran and O'Donoghue (2002) report that in 2000, London prices were, on average, 6.8 % more expensive; goods in London were on average 2.6% more expensive and services were 13.0 % more expensive. Due to data limitation, the disaggregation of rents by education has not been feasible, however this could have helped more in explaining the different cost-of-living faced by workers with different levels of education, located in different British regions.

This confirms that a more appropriate RPI should be constructed based on the regional figures rather than the national one.

The estimates reported in columns 4 and 5 include region fixed effects to control for unobserved regional heterogeneity. The estimates on the college-year interaction terms are not statistically different from the specifications that exclude regional fixed effects. The table shows that when using the regional real wage as the dependent variable, the conditional real wage difference between graduate and high school is smaller in real terms than in nominal; however this difference is not significantly different from zero.

As discussed by Moretti (2010), there are at least two aspects that might bias the estimates of the return to education and related wage differentials. The first concern might be related to unobserved differences in worker quality. The unobserved ability of graduates and high school graduates may vary differentially across regions and this could bias the estimates of the conditional wage differences between graduates and high school. Specifically what may be more important is the change over time in the average ability of college graduates relative to high school graduates in

a given region is systematically related to changes over time in housing prices in that region. In particular if average unobserved ability of graduates relative to high school graduates grows more (less) in expensive regions compared to less expensive regions, then the real graduates returns are biased downward (upward) (Moretti, 2010). Similarly Duranton and Monastiriotis (2002) suggest that the unobserved ability component that usually is included when measuring return to education, will not matter provided that there is no spatial bias in the distribution of unobserved abilities¹⁸. They argue that the most likely spatial selection is probably to be about higher unobserved abilities in London, any failure to correct for this when London stands out as being “more expensive” may lead to overestimates of the true regional inequalities.

The second element that might be a source for bias of the estimates relates to the unmeasured quality differences in housing; in fact the different cost-of-living faced by workers of different levels of education could also reflect differential changes in quality of housing (Moretti, 2010); for example the relative increase in the cost of housing experienced by college graduates may be overestimated if apartments rented by graduates are subject to more quality improvements than apartments in regions with many high school graduates. If these features have improved more in cities with many graduates, the estimates may be overestimating the relative increase in cost-of-living experienced by college graduates. The lack of attention towards the quality change of goods in any measure of inflation is a well known bias in the cost-of-living literature (Diewert, 1993). Hausman (2002) explains that the “constant basket“ approach ignores, among other aspects, the quality change in existing goods and that a use of a cost-of-living index based on utility (or expenditure functions) allows estimation of each of the effects of substitution, new goods and quality change. To estimate these effects, both price and quantity data are needed, unfortunately the latter are usually not available to the researcher.

Following Moretti, Table 13 restricts the analysis to UK born only workers though they are included in the additional specifications; excluding immigrants in the baseline regressions can be motivated by fact that in the US context immigration is often viewed as a proximate cause of the rising wage gap between high and low skilled workers (Card, 2010). Though the skills composition of migrants in the UK is different from that of the US (Rienzo, 2010), Manacorda et al. (2007) provide evidence that the native-immigrants wage differential is sensitive to the share of immigrants

¹⁸ Duranton and Monastiriotis (2002) explain that this bias selection problem can take three forms: First unobserved regional fixed-effects could lead to different educational choices for youngsters of similar abilities (or different participation choices for females); a second type of bias could be due to the migration patterns leading to an uneven spatial distribution of unobserved abilities; third unobserved ability may affect the probability of being in full-time employment. For example if the probability to find a full time job differs across regions, the distribution of unobserved characteristics for individuals at work across regions will be different.

in the working age population; in fact they show that a 10% rise in the population share of immigrants is estimated to increase native-migrant wage differential by 2%.

Immigrants, defined as workers born outside the UK, represent an increasing part of the working population (about 14% in 2009, Wadsworth 2010). In addition, the share of immigrants with a graduate level of education has been increasing over time. Another relevant reason to include immigrants in the estimations is motivated by the fact that the concentration of immigrants in London and the South-East has been increasing over time, for example in 2009 about 39% of the London population was made up of foreign-born people (Rienzo, 2010). Table 14 extends the analysis to immigrants. The returns to education are now lower (columns 1 and 2) than when immigrants were excluded, the nominal wage gap in 1997 was about 2% higher than the real one in the same year and 5% higher in 2008. The decrease in both the nominal and real conditional difference between wage of graduate workers and high school is higher (respectively 0.025 and 0.038) but still not statistically significant.

Table 15 reports similar estimates separately for men and women aged 25 to 60, only UK born, working both part time and full time. The returns to graduate education are higher for women than men in any time period. Deflating by the regional RPI makes little difference to the estimated graduate returns for both men and women in any year. Over time, although the conditional wage difference between college and high school increases for men, both in nominal and real terms, and decreases more for women, there is no statistically significant change in the graduate returns over time.

Table 16 reports similar estimates to tables 14 and 15 but extends the sample to cover the working age population (16 to 59 for women and 16 to 64 for men). The table documents that the estimated graduate to high school are sensitive to the sample age selection. Panel A of table 16 reports results from similar models to those in table 14. Considering all workers (men and women together), the conditional nominal wage differential between graduate and high school workers is now higher though not significantly (-0.033) than when restricting the analysis to workers aged 25-60 only (0.025); when using the regional deflator the decrease is now 0.045 compared to 0.038 when excluding younger workers. Panel B and C extend the analysis to men and women separately. The magnitude of both nominal and real conditional wage differentials is lower for men. For women, both the nominal and real conditional wage differentials are higher than men. The increases in the conditional (nominal and real) wage differentials when including younger workers may be due to cohort effects since the tenure of workers may change by education and age. This implies that cohorts of different levels of education may be affected differently by the cost of housing.

However, it has not been possible to analysis the effects of cost of housing by cohorts because longitudinal data would be more appropriate to that purpose.

Table 17 expands the estimates by analysing the wage differential between graduate and less than high school workers; this may be a proxy for the 90-10 gap which, over the sample period analysed, has been higher than any other measures of wage dispersion. The estimates in Table 17 are based on a sample of workers aged 25-60, UK born only, working part time and full time. When using the Regional RPI the wage gap is lower than the nominal conditional one, though not significantly, and this is true in all years reported (1997, 2002 and 2008). In this case the nominal and real wage differentials are always decreasing either when considering the whole sample (men and women together) or men and women separately. In all cases the regional deflator decreases the wage differentials; the magnitude is higher for women (0.128 and 0.143 respectively nominal and real) than it is for men (0.070 and 0.092 respectively nominal and real).

Including workers with lower levels of educational results appears to influence the sample included by increasing the magnitude of both nominal and real conditional wage differentials.

6. Conclusion

The existing literature investigating the trends in and causes of wage inequality in the UK usually measures wages in real terms by deflating nominal wages using the national Retail Prices Index (RPI). However the RPI does not account for differences in regional housing costs. Expenditure on housing is the largest component of total household expenditure and varies considerably through regions in the UK. Over time, housing costs have grown differentially across regions. Moreover, graduate workers appear to be more concentrated in more expensive British regions and increasingly so over time.

This paper has shown that when accounting for regional differences in the cost of housing the most common measure of UK inflation, the RPI, appears not to fully represent the cost-of-living in the various British regions. The national RPI underestimates the cost-of-living of workers living in the regions with the most expensive housing (London and the South East) and overestimates the cost-of-living for “cheaper” housing regions (Northern Ireland, Scotland). This inevitably has some implications when using the National or regional RPI to deflate the hourly wage.

When deflating hourly wages by the regional RPI, the average level of wages is lower by 8% to 11% an hour for all workers in London and the South East, whilst it is higher by 2% to 9% in

the remaining regions. However, though the use of a regional deflator makes a significant difference to levels, it does not make much difference to the graduate high school wage gap in any year or over time.

This paper shows how a regional deflator could be used in principle and further work could be based on extending this. The use of deflators and their measures are crucial in terms of policy decisions: since the decisions about how we compute inflation statistics can have a direct impact on policy decisions (Checchetti, 2007). Acknowledging the regional disparities in the cost-of-living in the UK also means that a study of regional variations in the cost-of-living has several important implications; Borooah et al (1996) pointed out three: first, there is the adjustment of social security benefit levels to take account of regional differences in prices. Secondly, conclusions about the relative deprivation or prosperity of regions, as measured by real disposable income, could also be susceptible to change in the face of regional variations in the cost-of-living. Lastly, conclusions about the number of persons living in poverty could also alter when regional cost-of-living variations are allowed for. Moreover, future research should also look at how differences in the regional cost-of-living should be taken into account to set minimum wages at a regional basis rather than at a national one.

As pointed out by Meullbauer and Murphy (2008), housing, location and demographic choices are closely connected. Housing markets are crucial for understanding regional evolutions and regional disparities in economic activities and living standards. Moreover, migration between regions plays a role in the working of regional housing and labour markets. House prices and the related cost-of-living have several effects on the labour-market and on the choices made by households for household formation and location.

The failure of the National RPI to appropriately reflect the real cost-of-living of different UK regions suggests the need for regional specific studies and related policy to address the existing regional differences in the labour market and standards of living; the persistence in regional unemployment rates is perhaps one symptom of those differences. The attention to more regional oriented analysis is also motivated by the fact that London, for example, is the most unequal region in the UK; although it has the highest proportion of households in the top tenth of income nationally, it also has the highest rate of income poverty of any region in England, with the highest proportion of people of all ages living below the poverty rate (The Guardian, 2009).

References

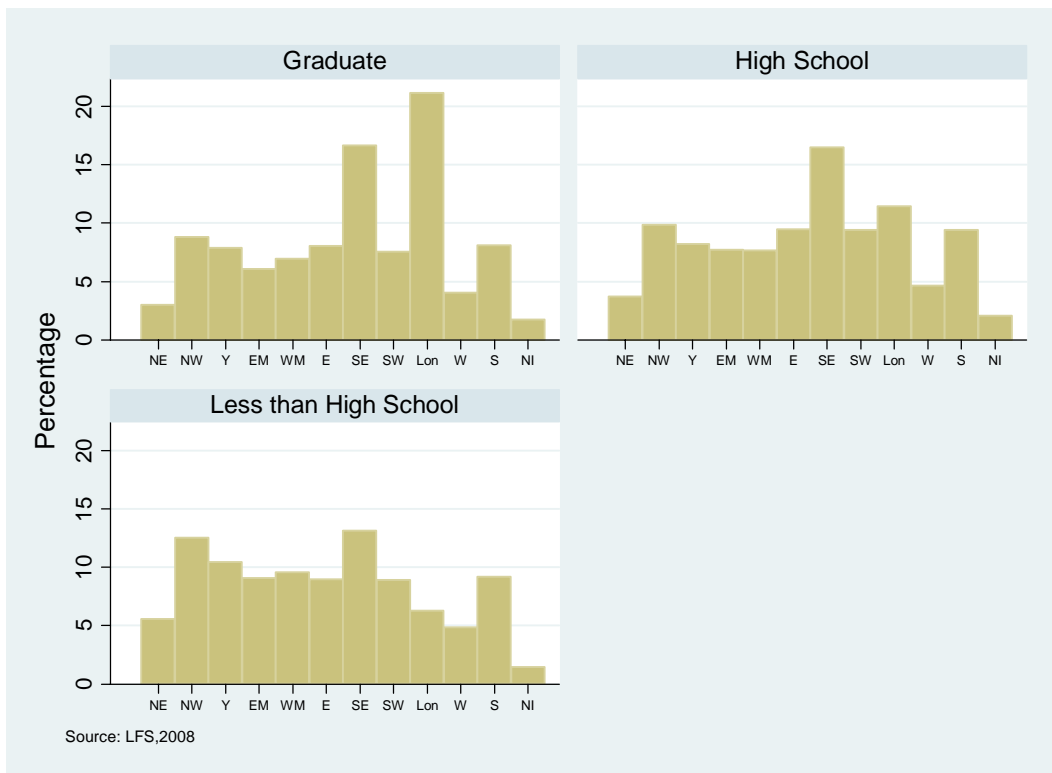
- Autor D., L.F. Katz, and M. Kearney (2005) 'Rising Wage Inequality: the Role of Composition and Prices', NBER Working Paper No. 11628.
- Baran, D. (2004) 'Consumer prices index and retail prices index: updating weights for 2004.' ONS.
- Baran, D. and J. O' Donoghue (2002) 'Price levels in 2000 for London and the regions compared with the national average', ONS.
- Bell, B and J. Van Reenen (2010) 'Bankers' pay and extreme wage inequality in the UK', CEP DP21.
- Bernard, A. B., Redding, S. J. , Schott P. K. and H. Simpson (2003) 'Relative wage variation and industry location', Oxford Bulletin of Economics and Statistics, Vol.70, Issue 4, p. 431-459.
- Black, D., Kelensikova, N., and L.J. Taylor (2009) 'Earnings function when wages and prices vary by location', Journal of Labour Economics.
- Blanchflower, D., and A. Oswald (2003) 'Calculating an appropriate regional funding adjustment for Worcestershire: a report for Worcestershire County Council', Worcestershire County Council.
- Blanchflower, D., and A. Oswald (2005) 'Regional wages and the need of a better area cost adjustment', Public Money and Management, Vol. 25 (2).
- Blow, L., and I. Crawford (2001) 'The cost of living with the RPI: Substitution bias in the UK retail prices index', The Economic Journal 111 (472, Features) (Jun.): F357-82.
- Borooah V. K., McGregor P.P.L., McKee P.M., and G.E. Mulholland (1996) 'Cost-of-living differences between the regions of the United Kingdom', in *New Inequalities, The changing distribution of income and wealth in the United Kingdom.* (Eds.) J. Hills.
- Card, D., and T. Lemieux (2001) 'Can falling supply explain the rising return to college for younger men? A cohort-based analysis', The Quarterly Journal of Economics 116, (2) (May):705-46.
- Checchetti, S. (2007) Housing in inflation measurement. VOX
- Consumer price indices technical manual. (2007) *Office for National Statistics*.
- Crawford I. (1996) 'UK household cost-of-living indices, 1979-92', in *New Inequalities, The changing distribution of income and wealth in the United Kingdom.*(Eds.) J. Hills.
- Crawford, I., and Z. Smith (2002) 'Distributional aspects of inflation', *Institute of Fiscal Studies* Commentary 90.

- Deaton, A. and J. Muellbauer (1980a) *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.
- Diewert, E. (1993) 'The early history of Price Index research', in *Essays in Index number theory*, Volume 1. W.E. Diewert and A.O. Nakamura, eds. Amsterdam, Elsevier, pp33-65.
- Diewert, E. (2003) 'The treatment of owner occupied housing and other durables in a consumer price index', Discussion paper No. 03-08, University of British Columbia.
- Duranton G., and V. Monastiriotis (2002) 'Mind the gap: The evolution of regional earnings inequalities in the U.K., 1982-1997', *Journal of Regional Science* 42 (1).
- Fenwick, D, and J. O' Donoghue (2004) 'Developing estimates of relative regional consumer price levels', ONS.
- Gosling, A. and T. Lemieux (2004) Labor Market Reforms and Changes in Wage Inequality in the United Kingdom and the United States In: *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000* eds. Card, D., Blundell, R. and Freeman, B.R.
- Green, F., and Y. Zhu (2008) 'Overqualification, job satisfaction, and increasing returns to graduate education', *Oxford Economic Papers*, 1-24.
- Hausman, J. (2003) 'Sources of bias and solutions to bias in the Consumer Price Index', *The Journal of Economic Perspective*, Vol. 17 N.1: 23-44.
- Hayes, P. (2005) 'Estimating UK regional price indices 1974-96', *Regional Studies* 39 (3) (05): 333-44.
- Henley, A. (2005) 'On regional growth convergence in Great Britain', *Regional Studies* 39 (9) (12): 1245-60.
- Holly, S., M. H. Pesaran, and T. Yamagata (2010) 'Spatial and temporal of house prices in the UK', *IZA Discussion Paper* DP 4694.
- Inman, P. (2010) 'Inflation index move is 'insidious' benefit cut, say TUC' from *The Guardian*, 12 October.
- Konus A.A. (1924) 'The problem of the true index of the cost of living', Reprinted in *Econometrica*, vol.7 page 10-29.
- Lemieux, T. (2006) 'Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?', *The American Economic Review* 96, (3) (Jun.): 461-98.
- Leunig, T., and H. Overman (2008) 'Spatial patterns of development and the British Housing Market', *Oxford Review of Economic Policy*. Vol. 24 (1): 59-78.

- Machin, S. (1998) 'Wage Inequality in the 1970s, 1980s and 1990s', In P.Gregg and J. Wadsworth (eds) *The State of Working Britain*, Manchester University Press.
- Machin, S. 2003. 'Wage inequality since 1975', in R. Dickens, P. Gregg, and J. Wadsworth (eds) *The Labour Market under new Labour*, Palgrave Macmillan, Basingstoke.
- Machin, S., and J. Von Reenen (2008) 'Changes in wage inequality', New Palgrave Dictionary of Economics.
- Manacorda, M., A., Manning, and J. Wadsworth (2007) 'The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain', Research in Labour Economics, 26, 125-155.
- Moretti, E. (2010) 'Real wage inequality', *NBER Working Paper no. 14370*.
- Muellbauer, J., and A. Murphy (2008) 'Housing markets and the economy: the assessment', Oxford Review of Economic Policy, Vol. 24 (1): 1-33.
- Nickell, S. (2005) 'Practical issues in UK monetary policy, 2000-2005, Bank of England.
- Office for National Statistics*. (2007) Consumer price indices technical manual. 2007. Technical Report.
- Office for National Statistics*. (2009) Family Spending. A report on the 2008 Living costs and Food Survey. (Eds.) Skentelbery R.
- Rienzo, C. (2010) 'Increasing wage inequality: a comparative analysis between Italy, UK and the US', PhD Chapter.
- Rienzo, C. (2010) 'Residual wage inequality and immigration in the UK and the U.S.', PhD Chapter.
- Rienzo, C. (2010) 'Stock of migrants in the UK', report for the Migration Observatory, University of Oxford.
- Schmitt, J. (1995) 'The changing Structure of Male earnings in Britain, 1974-1988', in R. Freeman and L.Katz (eds.), *Changes and Differences in Wage Structures* Chicago: University of Chicago Press.
- Sloane, P.J. (2003) 'Much ado about nothing? What does the overeducation literature really tell us?', In F. Buchel, A. de Grip, and A. Martens (eds) *Overeducation in Europe: current issues in Theory and Policy*, Edward Elgar, Cheltenham.
- Wadsworth, J. (2010) 'Immigration and the UK Labour Market: evidence from economic research', CEP discussion paper EA006.
- Walker, I. and Y. Zhu (2008) 'The college wage premium and the expansion of higher education in the UK', Scandinavian Journal of Economics, 110, 695-709.

Wooldridge, J. (2006) Introductory econometrics a modern approach.

Figure1: Distribution of Workers by Education Within Region, All Workers, 2008



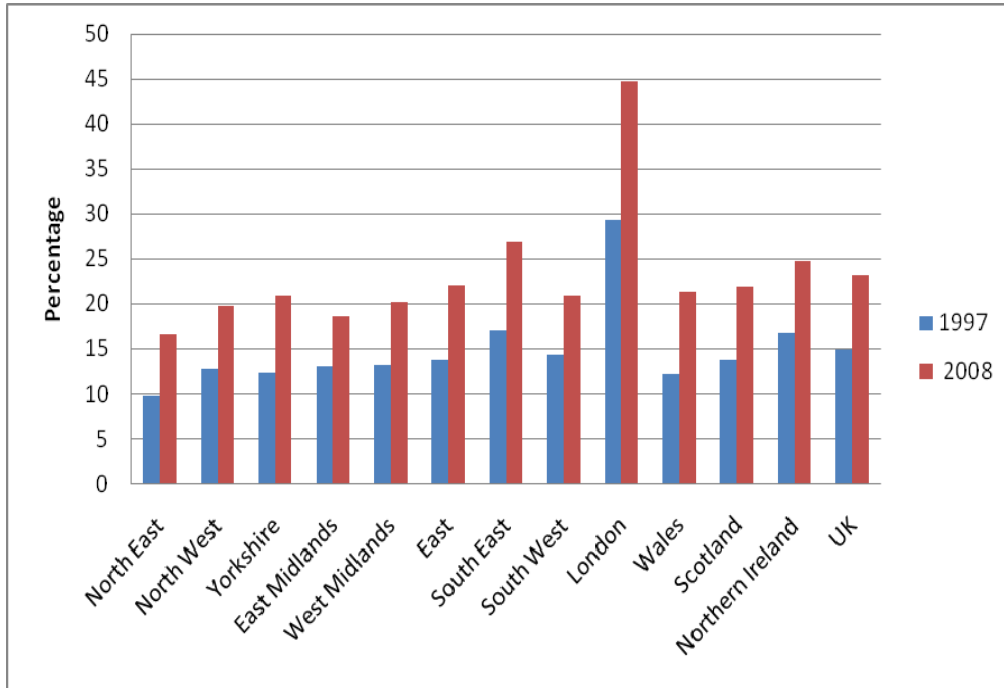
Notes: NE= North East; NW=North West; Y=Yorkshire; EM=East Midlands; WM=West Midlands; E=East; SE=South East; SW=South West; Lon=London; W=Wales; S=Scotland; NI=Northern Ireland. Sample is based on men (16-64) and women (16-59) working full time and part time, employees and main job only. Graduate refers to anyone who left full time education at 21 or later; “High school graduate” refers to anyone who left full time education between the ages of 17 and 20; “less than high school” refers to those who left education at age 16 or less.

Table 1: Distribution of Workers by Education across Regions, 2008

	Graduate	High School	Less than High School	% of total working age population
North East	3.0	3.7	5.6	4.4
North West	8.8	9.8	12.5	10.8
Yorkshire	7.9	8.2	10.5	9.1
East Midlands	6.1	7.7	9.1	7.9
West Midlands	6.9	7.7	9.6	8.3
East	8.1	9.5	9.0	8.9
South East	16.7	16.5	13.1	15.0
South West	7.5	9.4	8.9	8.7
London	21.1	11.4	6.2	11.5
Wales	4.0	4.7	4.9	4.6
Scotland	8.1	9.4	9.2	9.0
Northern Ireland	1.8	2.1	1.5	1.7
Total	100	100	100	100

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Graduate refers to anyone who left full time education at 21 or later; “High school graduate” refers to anyone who left full time education between the ages of 17 and 20; “less than high school” refers to those who left education at age 16 or less.

Figure 2:
 Percentage of Graduate Workers by Region, 1997 and 2008



Notes: Based on LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Graduate refers to anyone who left full time education at 21 or later.

Table 2: Distribution of Workers by Education, UK Regions 1997-2008

Region	1997			2008		
	Graduate	High School	Less than High School	Graduate	High School	Less than High School
North East	10.3	19.5	70.2	16.6	26.5	56.8
North West	13.1	22.2	65.1	19.4	28.5	51.7
Yorkshire	12.7	21.9	65.7	20.8	28.0	51.0
East Midlands	13.5	24.6	62.3	18.7	30.3	51.0
West Midlands	13.6	23.4	63.4	20.1	28.6	51.2
East	14.1	28.4	57.9	21.5	33.1	44.9
South East	17.5	31.6	51.3	26.7	34.3	38.8
South West	14.6	29.8	55.9	20.8	33.6	45.5
London	28.9	33.2	37.5	42.9	31.5	25.1
Wales	12.2	28.0	59.8	21.5	31.6	47.1
Scotland	14.0	26.4	59.8	21.7	32.6	45.5
Northern Ireland	17.2	30.1	53.1	24.8	37.6	37.8

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Graduate refers to anyone who left full time education at 21 or later; "High school graduate" refers to anyone who left full time education between the ages of 17 and 20; "less than high school" refers to those who left education at age 16 or less.

Table 3: Price for Rent by Regions, 1997-2008

Region	1997	2008	1997-2008 weekly rent change	% change to the UK increase
North East	42.50	71.78	29.28	-28.1
North West	48.86	83.53	34.67	-14.9
Yorkshire & The Humber	43.20	78.64	35.44	-13.0
East Midlands	44.58	81.31	36.73	-9.8
West Midlands	48.46	84.39	35.93	-11.8
East	56.89	100.38	43.50	6.8
South East	69.87	123.88	54.01	32.5
South West	56.28	103.17	46.89	15.1
London	86.10	163.22	77.12	89.3
Wales	46.41	77.37	30.96	-24.0
Scotland	37.96	69.68	31.72	-22.1
Northern Ireland	35.93	68.63	32.70	-19.7
UK	51.42	92.17	40.75	

Source: Based on Family Resources Survey.

Table 4: Mean of Hourly Wage

Year		Nominal Wage	Real wage (National RPI)	Real Wage (Regional RPI)
1997	All workers	7.65	7.65	7.59
	Men	8.69	8.69	8.61
	Women	6.51	6.51	6.49
2000	All workers	8.86	8.42	8.31
	Men	9.97	9.47	9.35
	Women	7.63	7.25	7.16
2003	All workers	10.05	8.85	8.95
	Men	11.2	9.86	9.96
	Women	8.8	7.75	7.84
2006	All workers	11.33	9.3	9.57
	Men	12.48	10.24	10.53
	Women	10.07	8.27	8.54
2008	All workers	12.15	9.25	9.52
	Men	13.39	10.19	10.48
	Women	10.79	8.22	8.47

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level.

Table 5: Real Hourly Wage by Aggregate Regions, 1997-2008

Year	UK		London		South East		Rest of the UK	
	N. RPI	R. RPI	N.RPI	R. RPI	N. RPI	R. RPI	N. RPI	R. RPI
1997	7.65	7.58	9.58	8.47	8.58	8.17	7.17	7.33
1998	7.87	7.76	9.97	8.79	8.93	8.36	7.32	7.47
1999	8.09	8.00	10.23	9.13	9.10	8.59	7.53	7.69
2000	8.42	8.31	11.01	9.60	9.41	8.96	7.80	7.97
2001	8.74	8.72	11.59	10.13	9.71	9.18	8.07	8.39
2002	8.8	8.86	11.46	10.16	9.89	9.46	8.15	8.52
2003	8.85	8.95	11.63	9.90	9.97	9.45	8.19	8.70
2004	9.02	9.17	11.66	9.87	9.97	9.47	8.41	9.00
2005	9.20	9.43	11.99	10.27	10.11	9.67	8.58	9.26
2006	9.30	9.57	11.92	10.41	10.20	9.79	8.70	9.39
2007	9.33	9.63	11.96	10.13	10.37	10.15	8.71	9.45
2008	9.25	9.52	12.02	9.84	10.22	9.78	8.60	9.41

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. N. refers to National, R: refers to regional.

Table 6: Measures of Wage Dispersion and Changes over Time, Log Hourly Wage

Year	All workers	Real wage RPI (National)	Real Wage RPI (Regional)	Difference
Panel B:		Wage Dispersion		
1997	Standard Dev.	0.588	0.582	-0.006
	Variance	0.346	0.339	-0.008
	90-50	0.766	0.750	-0.016
	90-10	1.396	1.375	-0.021
	50-10	0.630	0.625	-0.005
2000	Standard Dev.	0.580	0.572	-0.009
	Variance	0.337	0.327	-0.010
	90-50	0.760	0.749	-0.011
	90-10	1.370	1.351	-0.019
	50-10	0.611	0.602	-0.009
2003	Standard Dev.	0.557	0.546	-0.011
	Variance	0.310	0.298	-0.012
	90-50	0.768	0.749	-0.019
	90-10	1.351	1.320	-0.031
	50-10	0.583	0.571	-0.012
2006	Standard Dev.	0.561	0.553	-0.009
	Variance	0.315	0.306	-0.010
	90-50	0.763	0.752	-0.011
	90-10	1.349	1.316	-0.033
	50-10	0.585	0.564	-0.021
2008	Standard Dev.	0.567	0.557	-0.010
	Variance	0.321	0.310	-0.011
	90-50	0.767	0.751	-0.016
	90-10	1.347	1.326	-0.021
	50-10	0.579	0.575	-0.005
Panel B:		1997-2008 Change in Wage Dispersion		
		National RPI	Regional RPI	
	Standard dev.	-0.021	-0.025	-0.004
	Variance	-0.025	-0.029	-0.004
	90-50	0.001	0.001	0.000
	90-10	-0.049	-0.049	0.000
	50-10	-0.051	-0.050	0.001

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level.

Table 7: Measures of Wage Dispersion, by Region 2008

	Standard Dev.	90-10	90-50	50-10
North East	0.508	1.169	0.68	0.489
North West	0.518	1.249	0.708	0.541
Yorkshire	0.518	1.229	0.705	0.524
East Midlands	0.551	1.267	0.752	0.515
West Midlands	0.543	1.291	0.749	0.541
East	0.579	1.388	0.79	0.597
South East	0.590	1.413	0.788	0.625
South West	0.555	1.271	0.722	0.549
London	0.602	1.508	0.781	0.728
Wales	0.534	1.272	0.764	0.508
Scotland	0.541	1.304	0.753	0.551
Northern Ireland	0.502	1.169	0.703	0.466
UK	0.557	1.326	0.751	0.575

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Using regional RPI.

Table 8: Changes in Real Hourly Wage by Region, 1997-2008

Region	Based on National RPI (1)	Based on Regional RPI (2)
North East	1.18	2.10
North West	1.28	1.99
Yorkshire	1.37	2.07
East Midlands	1.15	1.78
West Midlands	1.47	2.13
East	1.54	1.99
South East	1.64	1.61
South West	1.50	1.80
London	2.43	1.37
Wales	1.54	2.32
Scotland	1.65	2.54
Northern Ireland	1.74	2.39
UK	1.59	1.93

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level.

Table 9: Changes in Weekly Rent, by Education Group

	1997	London	South East
	UK		
Graduates	98.31	112.40	118.62
High School	65.77	85.27	78.59
Less than high School	46.30	61.03	58.25
	2008		
Graduates	164.39	244.28	187.44
High School	105.72	160.84	134.61
Less than high School	84.58	112.32	101.21
	1997-2008	Percentage	Increase
Graduates	67.2%	117.3%	58.0%
High School	60.7%	88.6%	71.3%
Less than high School	82.7%	84.1%	73.7%

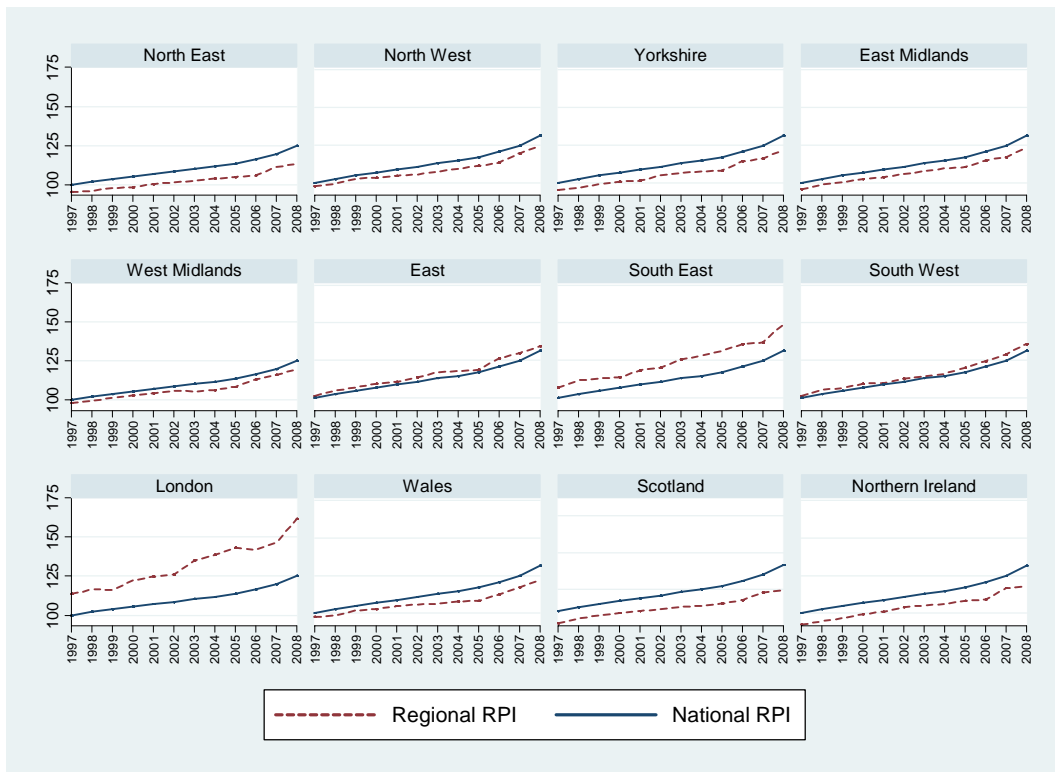
Based on the FRS.

Table 10: Real Hourly Wage for Graduate Workers and High School Workers

	1997	2002	2008	1997-2008 change	1997	2002	2008	1997-2008 change
Panel A								
Graduates								
	Men				Women			
National RPI	12.95	14.61	14.08	1.13	10.31	11.72	11.6	1.28
Regional RPI	12.63	14.41	14.02	1.38	10.1	11.6	11.67	1.57
Diff.	0.32	0.2	0.07	-0.25	0.21	0.11	-0.07	-0.28
Percentage	2.5%	1.4%	0.5%	-18.2%	2.1%	1.0%	-0.6%	-18.1%
Panel B								
High School graduates								
	Men				Women			
National RPI	9.25	10.09	9.93	0.68	6.88	7.7	8.01	1.14
Regional RPI	9.11	10.1	10.22	1.11	6.78	7.75	8.28	1.49
Diff.	0.14	-0.01	-0.29	-0.43	0.09	-0.05	-0.26	-0.36
Percentage	1.5%	-0.1%	-2.9%	-38.9%	1.4%	-0.6%	-3.2%	-23.9%

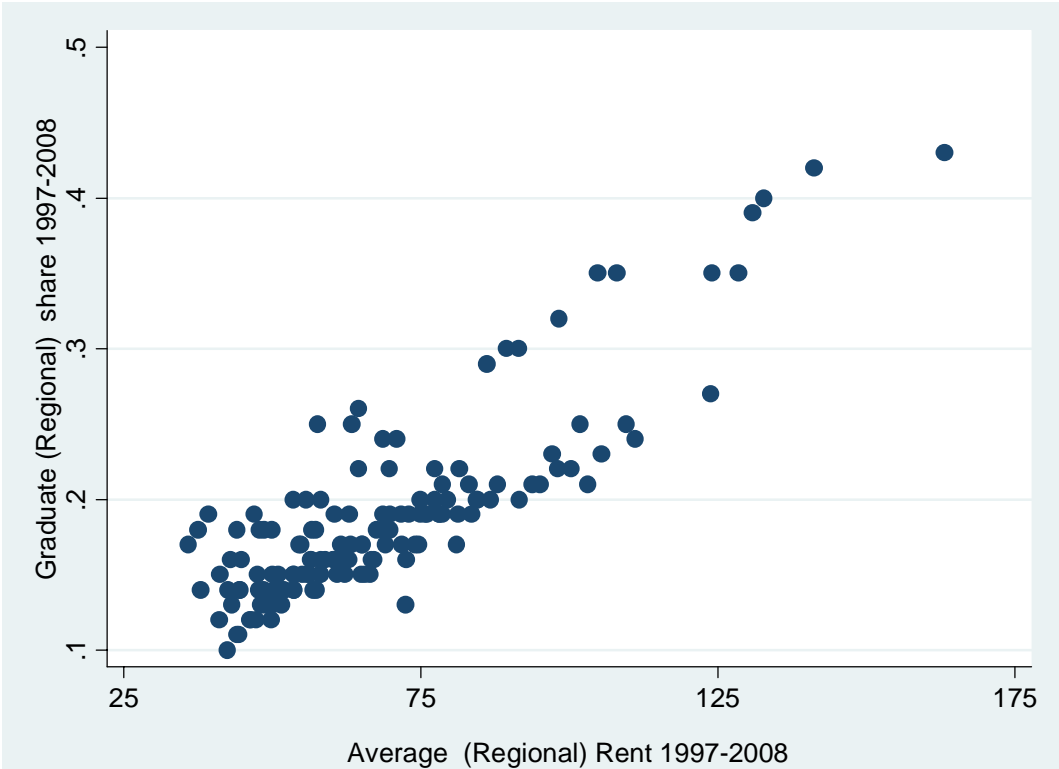
Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Graduate refers to anyone who left full time education at 21 or later; "High school graduate" refers to anyone who left full time education between the ages of 17 and 20.

Figure 3: National and Regional RPI, UK 1997-2008



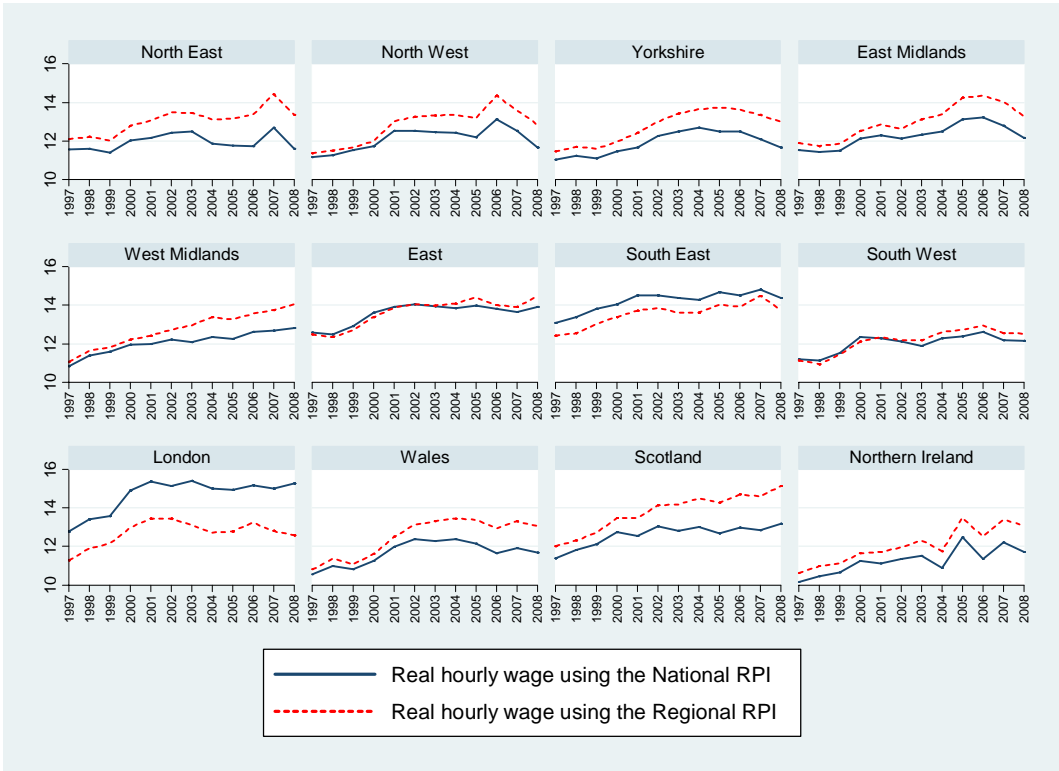
Source: EFS/FES/LCS and FRS.

Figure 4: How Increasing Share of Graduates Relate to Increasing Price for Rent



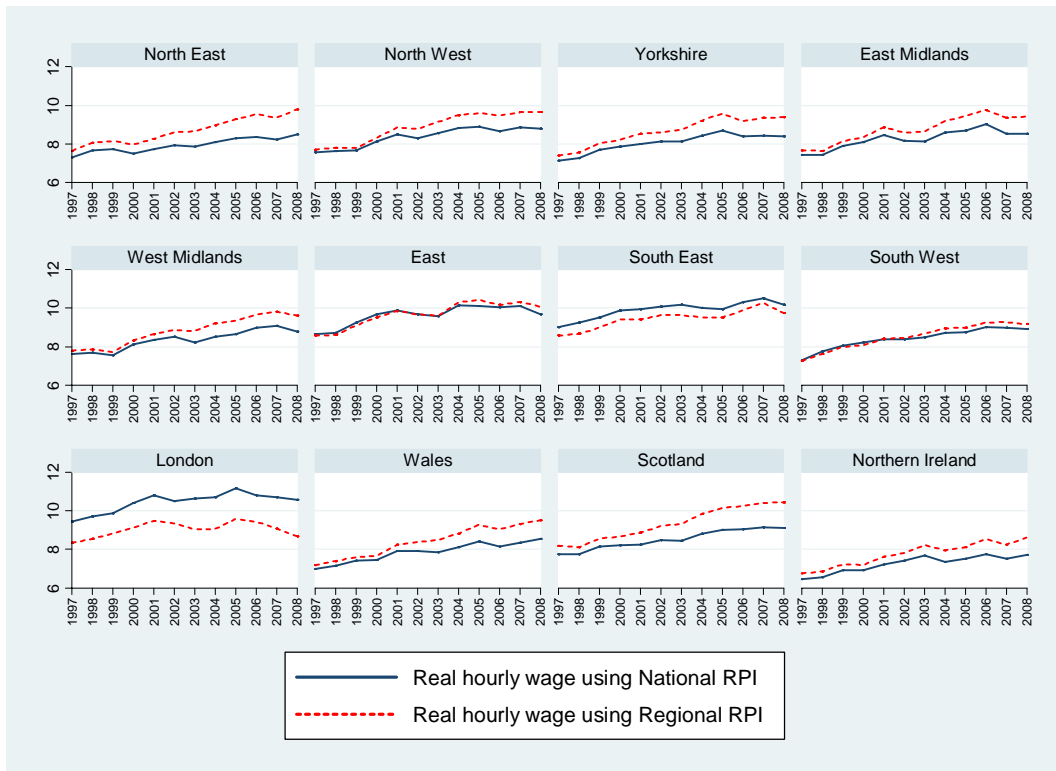
Notes: Based on the FRS and LFS, 1997-2008.

Figure 5a: Real Hourly Wage for Graduate Workers by Region,
All Workers 1997-2008



Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Graduate refers to anyone who left full time education at 21 or later.

Figure 5b: Real Hourly Wage for High School Workers by Region, All Workers 1997-2008



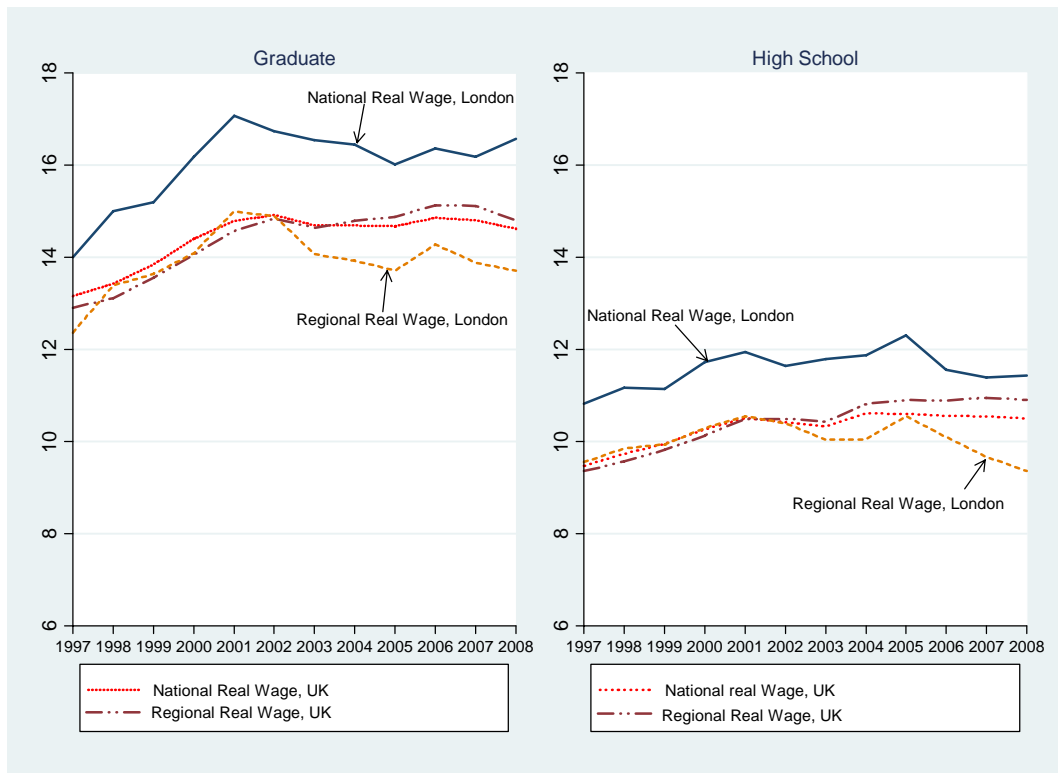
Notes: Based on LFS from 1997-2008; sample includes all workers in labour force, employed and main job only. High School workers are defined as those who left school between 17 and 20 years old.

Table 11: Real Hourly Wage for Graduate Workers by Region, 2008

Region	based on National RPI	based on Regional RPI	Difference
	(1)	(2)	(3)
North East	11.23	12.95	1.72
North West	11.32	12.46	1.14
Yorkshire	11.27	12.59	1.32
East Midlands	11.79	13.01	1.22
West Midlands	12.43	13.61	1.18
East	13.27	13.80	0.53
South East	13.83	13.23	-0.60
South West	11.79	12.14	0.35
London	14.88	12.18	-2.70
Wales	11.33	12.64	1.31
Scotland	12.73	14.62	1.89
Northern Ireland	11.27	12.61	1.34
UK	12.93	12.92	-0.01

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Graduate refers to anyone who left full time education at 21 or later.

Figure 6a: Real Hourly Wage for Graduate and High School Male Workers, London and UK 1997-2008



Notes: Based on the LFS from 1997-2008; sample includes males aged 16-64, employed, full time and part time, and main job only. Graduate workers are defined as those who left school at age 21 or more. High school workers are defined as those who left school between 17 and 20 years old.

Figure 6b: Real Hourly Wage for Graduate and High School Female Workers, London and UK 1997-2008



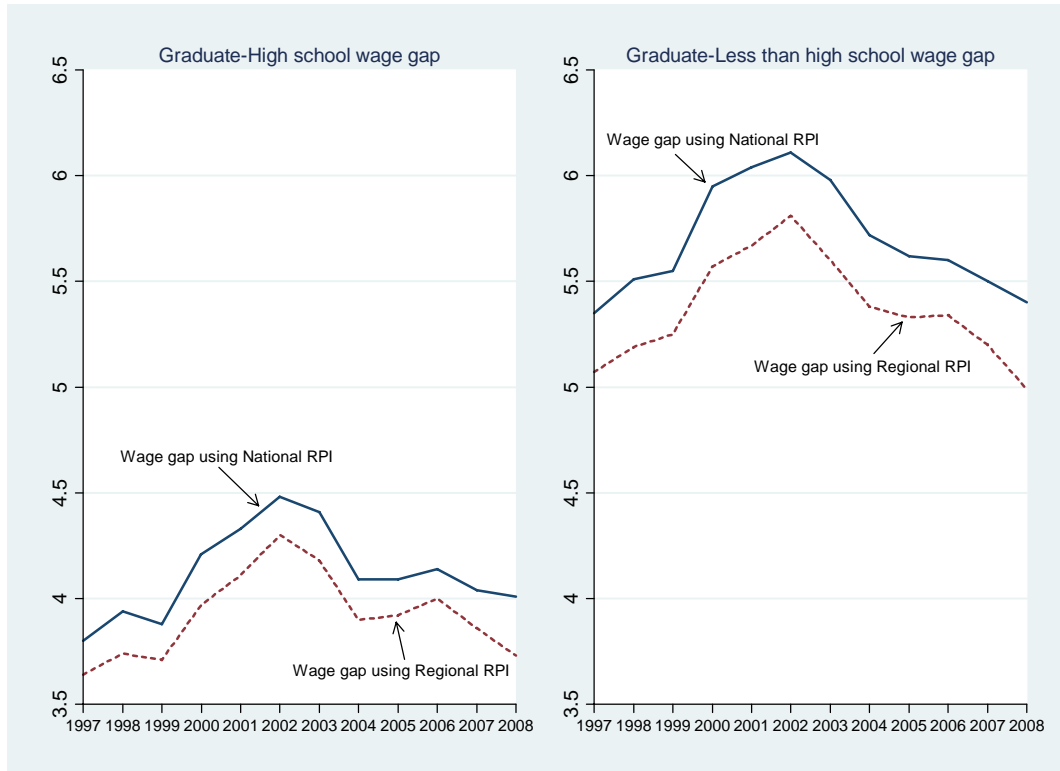
Notes: Based on LFS from 1997-2008; sample includes female workers aged 16-59, employed, full time and part time and main job only. Graduate workers are defined as those who left school at age 21 or more. High school workers are defined as those who left school between 17 and 20 years old.

Table 12: Graduate-high School and Graduate-less than High School

Year	Grad.-High School Wage Gap			Grad.-less than High School Wage Gap			
	National RPI	Regional	Difference	National	RPI	Regional RPI	Diff.
1997	3.80	3.64	0.16	5.35		5.07	0.28
1998	3.94	3.74	0.20	5.51		5.19	0.32
1999	3.88	3.71	0.17	5.55		5.25	0.30
2000	4.21	3.97	0.24	5.95		5.57	0.38
2001	4.33	4.11	0.22	6.04		5.67	0.37
2002	4.48	4.30	0.18	6.11		5.81	0.30
2003	4.41	4.18	0.23	5.98		5.60	0.38
2004	4.09	3.90	0.19	5.72		5.38	0.35
2005	4.09	3.92	0.17	5.62		5.33	0.30
2006	4.14	4.00	0.14	5.60		5.34	0.26
2007	4.04	3.86	0.18	5.50		5.20	0.30
2008	4.01	3.73	0.28	5.40		4.99	0.41
1997-2008 Change	0.21	0.09	0.12	0.05		-0.08	0.13

Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Graduate refers to anyone who left full time education at 21 or later.

Figure 7: Graduate-High School and Graduate-Less than High School Wage Gap, UK 1997-2008



Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Graduate refers to anyone who left full time education at 21 or later; “High school graduate” refers to those who left full time education between the age of 17 and 20.

Table 13: Nominal and Real Conditional Wage Difference between Workers with a High School Degree and Workers with College or more, UK Born

	1997	2008	1997-2008 Change	1997	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Model 1</u>						
Nominal Wage Difference	0.419*** (0.010)	0.402*** (0.017)	-0.017 (0.000)	0.408*** (0.008)	0.388*** (0.007)	-0.014 (0.000)
<u>Model 2</u>						
Real Wage difference- Regional RPI	0.413*** (0.009)	0.395*** (0.016)	-0.018 (0.000)	0.410*** (0.007)	0.382*** (0.012)	-0.028 (0.000)
Region Fixed Effects	No	No		Yes	Yes	

* p<0.05, ** p<0.01, *** p<0.001

Notes: Standard errors clustered by region in parentheses. Sample is based on men and women aged 25-60, employees, working full time and part time, main job only UK born only. The dependent variable in Model 1 is the log of nominal hourly wage. The dependent variable in Model 2 is the log of real hourly wage, deflated by regional RPI. Controls include a cubic in potential experience, year fixed effects, gender and race.

Table 14: Nominal and Real Conditional Wage Difference between Workers with a High School Degree and Workers with College or more, Migrants and UK Born

	1997	2008	1997-2008 Change (3)	1997	2008	1997-2008 Change (6)
	(1)	(2)		(4)	(5)	
<u>Model 1</u>						
Nominal Wage Difference	0.408*** (0.010)	0.383*** (0.014)	-0.025 (0.000)	0.399*** (0.012)	0.402*** (0.011)	0.003 (0.000)
<u>Model 2</u>						
Real Wage difference- Regional RPI	0.401*** (0.007)	0.363*** (0.010)	-0.038 (0.000)	0.371*** (0.011)	0.364*** (0.011)	-0.007 (0.000)
Region Fixed Effects	No	No		Yes	Yes	

Notes: Standard errors clustered by region in parentheses. Sample is based on men and women aged 25-60, employees, working full time and part time, main job. The dependent variable in Model 1 is the log of nominal hourly wage. The dependent variable in Model 2 is the log of real hourly wage, deflated by regional RPI. Controls include a cubic in potential experience, year fixed effects, gender and race.

Table 15: Nominal and Real Conditional Wage between Workers and with College or more and High School Degree, by Men and Women

	1997	2002	2008	1997-2008 Change	1987	2002	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Men								
Nominal Wage Difference	0.366*** (0.009)	0.394*** (0.014)	0.370*** (0.016)	0.004 (0.000)	0.357*** (0.006)	0.381*** (0.011)	0.354*** (0.010)	-0.003 (0.000)
Real Wage difference- Regional RPI	0.360*** (0.007)	0.386*** (0.011)	0.351*** (0.009)	-0.009 (0.000)	0.359*** (0.006)	0.382*** (0.011)	0.348*** (0.010)	-0.011 (0.000)
Panel B. Women								
Nominal Wage Difference	0.466*** (0.015)	0.471*** (0.016)	0.427*** (0.019)	-0.029 (0.000)	0.452*** (0.015)	0.457*** (0.012)	0.413*** (0.012)	-0.039 (0.000)
Real Wage difference- Regional RPI	0.460*** (0.014)	0.464*** (0.012)	0.411*** (0.015)	-0.049	0.454*** (0.014)	0.459*** (0.012)	0.407*** (0.015)	-0.047
Region Fixed Effects	No	No	No		Yes	Yes	Yes	

Notes: Standard errors clustered by region in parentheses. Sample includes workers aged 25-60, only UK born, working part-time and full-time. The dependent variable in Model 1 is the log of nominal hourly wage. The dependent variable in Model 2 is the log of real hourly wage, deflated by regional RPI. Controls include a cubic in potential experience, year fixed effects, gender and race.

Table 16: Additional Specifications: Nominal and Real Conditional Wage between Workers and with College or more and High School Degree: All Working Force Population

	1997	2002	2008	1997-2008 Change	1997	2002	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A.								
All Workers								
Nominal Wage Difference	0.430*** (0.009)	0.446** (0.014)	0.397*** (0.016)	-0.033 (0.000)	0.419*** (0.011)	0.432*** (0.010)	0.383*** (0.011)	-0.036 (0.000)
Real Wage diff. Regional RPI	0.421*** (0.011)	0.435*** (0.010)	0.376*** (0.009)	-0.045 (0.000)	0.422*** (0.010)	0.434*** (0.010)	0.376*** (0.010)	-0.046 (0.000)
Panel B. Men								
Nominal Wage Difference	0.389*** (0.007)	0.413*** (0.016)	0.380*** (0.020)	-0.009 (0.000)	0.381*** (0.008)	0.401*** (0.014)	0.365*** (0.014)	-0.016 (0.000)
Real Wage diff. Regional RPI	0.381*** (0.008)	0.402*** (0.011)	0.356*** (0.009)	-0.025 (0.000)	0.384*** (0.008)	0.403*** (0.013)	0.357*** (0.010)	-0.027 (0.000)
Panel C. Women								
Nominal Wage Difference	0.461*** (0.015)	0.468*** (0.015)	0.406*** (0.015)	-0.055 (0.000)	0.449*** (0.018)	0.452*** (0.012)	0.393*** (0.012)	-0.056 (0.000)
Real Wage diff. Regional RPI	0.453*** (0.017)	0.456*** (0.011)	0.388*** (0.012)	-0.065 (0.000)	0.451*** (0.017)	0.454*** (0.012)	0.386*** (0.013)	-0.065 (0.000)
Region Fixed Effects	No	No	No		Yes	Yes	Yes	

Notes: Standard errors clustered by region in parentheses Sample includes workers aged 16 to 64, employees, working full time and part time, main job only. The dependent variable in Model 1 is the log of nominal hourly wage. The dependent variable in Model 2 is the log of real hourly wage, deflated by regional RPI. Controls include a cubic in potential experience, year fixed effects, gender, race and control for migrants.

Table 17: Additional Specifications: Nominal and Conditional Wage Difference between Workers with a Graduate Degree or more and Workers with a Less than High School Graduates

	1997	2002	2008	1997-2008 Change	1997	2002	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. All workers								
<u>Model 1</u>								
Nominal Wage Difference	0.696*** (0.008)	0.689*** (0.015)	0.601*** (0.016)	-0.095 (0.000)	0.669*** (0.016)	0.663*** (0.012)	0.575*** (0.012)	-0.094 (0.000)
<u>Model 2</u>								
Real Wage difference- Regional RPI	0.675*** (0.013)	0.666*** (0.012)	0.562*** (0.017)	-1.113 (0.000)	0.677*** (0.013)	0.666*** (0.012)	0.562*** (0.017)	-0.115 (0.000)
Panel B. Men								
<u>Model 1</u>								
Nominal Wage Difference	0.650*** (0.014)	0.663*** (0.016)	0.580*** (0.018)	-0.070 (0.000)	0.623*** (0.017)	0.636*** (0.013)	0.551*** (0.013)	-0.072 (0.000)
<u>Model 2</u>								
Real Wage difference- Regional RPI	0.628*** (0.018)	0.640*** (0.013)	0.536*** (0.018)	-0.092 (0.000)	0.630*** (0.015)	0.640*** (0.012)	0.537*** (0.015)	-0.093 (0.000)
Panel C. Women								
<u>Model 1</u>								
Nominal Wage Difference	0.746*** (0.013)	0.713*** (0.020)	0.618*** (0.018)	-0.128 (0.000)	0.720*** (0.022)	0.688*** (0.018)	0.595*** (0.015)	-0.125 (0.000)
<u>Model 2</u>								
Real Wage difference- Regional RPI	0.727*** (0.017)	0.692*** (0.017)	0.584*** (0.017)	-0.143 (0.000)	0.726*** (0.019)	0.691*** (0.017)	0.584*** (0.017)	-0.142 (0.000)
Region Fixed Effects	No	No	No		Yes	Yes	Yes	

Notes: Standard errors clustered by region in parentheses Sample includes workers aged 25-60, employees, working full time and part time, main job only. The dependent variable in Model 1 is the log of nominal hourly wage. The dependent variable in Model 2 is the log of real hourly wage, deflated by regional RPI. Controls include a cubic in potential experience, year fixed effects, gender, race and control for migrants.