Essays on Wage Inequality:

the Role of Composition, Immigration and the Cost-of-living

Cinzia Rienzo

Department of Economics

Royal Holloway, University of London

Ph.D. Thesis

2011

Declaration of Authorship.

I Cinzia Rienzo hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed _____

Date_____

Abstract

This thesis focuses on the role of composition, immigration and the cost-of-living on wage inequality.

I begin by investigating to what extent changing characteristics of the labour force can help explain the fact that residual or within-group wage inequality –wage dispersion among workers with the same education and experience- is generally thought to account for most of the growth in wage inequality observed in several industrialised countries over the last thirty years. I compare the results for men and women in Italy, the UK, and the US from 1987 to 2003 or 2004. I find that even though residual does account for most of the wage variation in all countries, there is no common increasing trend in residual inequality. I also find that workforce composition does not always act to increase the residual wage inequality.

In the second part of the thesis, I investigate the effects of immigration on residual wage inequality in the UK and the US between 1994 and 2008, by assessing whether and to what degree immigration contributed, along with technology, institutions and traditional explanations, to widening inequality.

The analysis reveals that residual wage inequality is higher amongst immigrants than amongst natives. However, such differences do not contribute (much) to the increasing residual wage inequality observed in the two countries.

The final section of this thesis questions how existing estimates of inequality change when differences in the cost-of-living and the differential concentrations of individuals with different levels of education across regions are taken into account. I focus on changes in the difference in the hourly wage for workers with a college degree and high school degree in the UK between 1997 and 2008. 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, by 8% to 11% an hour for all workers in London and the South East, whilst it is higher, by around 2% to 9% in the remaining regions; similarly the level, but not changes, in wage inequality is lower when deflating by the real regional RPI.

Tables of contents

List of Tables

List of Figures

Acknowledgements

First, I would like to thank my supervisor Jonathan Wadsworth for his precious advice, several substantial suggestions and constant encouragement. Jonathan has been an invaluable mentor throughout all the stages of the PhD, allowing me the time to bring this thesis to fruition. He deeply influenced my understandings and research approach in Labour Economics.

During my study at Royal Holloway, I have also benefited from the comments and teaching of the Economics Department, especially Dan Anderberg, Arnaud Chevalier and Victor Lavy.

Numerous individuals have provided advice particularly David Card, Ian Crawford, Eric Gould, Jim O'Donoghue, Jennifer Hunt and Alberto Pozzolo.

I have also received several helpful comments from the seminar participants at the University of Molise, Royal Holloway, WPEG Annual Conference 2008, EALE 2008, Cemmap Conference at Birkbeck University; 3rd ICEEE (Italian Congress of Econometrics and Empirical Economics) at Università Politecnica delle Marche; PUEK PhD Workshop 2009 at IFS, IZA 12th Summer School, CEP (LSE) Labour Market Workshop, WPEG Annual Conference 2009 at the University of Nottingham and APPAM Conference 2010 at the University of Maastricht.

The British data used in this thesis were made available by the Data Archive; the Outgoing Rotation Group (ORG CPS) was made available through the Centre for Economic and Policy Research (CEPR) in Washington. My thanks go to the research division of the Bank of Italy for making the Survey for Household of Income and Wealth (SHIW) accessible. I am grateful to Thomas Lemieux for providing me with his data and for his valuable explanations and to Enrico Moretti for providing me with his dofile.

The completion of my studies would not have been possible without the financial support that I received from Royal Holloway.

Many friends have followed my PhD path; in particular I would like to thank Gaby for being an extraordinary friend and colleague, as well as Chiara for sharing with me all the worries and anxieties that a PhD brings.

My family has been an enormous support during the past years, helping me to get to the end despite adversities. I thank my mother Alfonsina, my brother Carmine and sister Mariantonietta. This thesis is dedicated to the memory of my father Antonio who could not see the end.

Finally, I would also like to thank Marco Manacorda for his guidance and for supporting me to embark on the PhD after my MSc.

Remaining errors are my own.

Chapter 1: Introduction

"…the distribution is a focal point at which the functioning of the economic system impinges upon the human beings […] .It follows that better knowledge and comprehension of the subject are indispensable, not only in and of itself but also as a step in learning more about the functioning of society-in both the long and short run."

Kuznetz, 1955

The study of the structure of wages has been a preoccupation of economists for a long time and dates back at least as far as Adam Smith. Since wages are an important part of household income and economic well-being, the trends (and increase) in the dispersion of wages mirror very closely the trends in income, consumption and poverty rates, affecting health condition as well as social exclusion.

The existence of and persistent increase in wage inequality has several important macroeconomic implications in an economy. Wage inequality is an indicator of the level of prosperity and performance of a country, since it directly reflects on the level and growth of the GDP, strongly influencing economic growth rates and convergence. Inequality in wage, implying inequality in income, is often associated with inequality in savings that determines the economic growth of a society (Kuznets 1955).

Economists pointing out the connections between health, earnings and education, have also emphasised the negative correlation between socio-economic status and various risky behaviours, such as smoking, excessive drinking, obesity, and lack of exercise.

Deaton (2000) gives evidence that health inequality in the United States is rising in tandem with income inequality. Growing income differences are associated both with an increase in spread in the distribution of health, as well as with an increase in the gradient linking education and health.

Wage differentials are often associated with labour market disparities (Blau and Kahn, 1996). This embodies a significant concern in many countries,

mainly because countries experiencing higher wage differentials and labour market disparities are also countries that differ dramatically in standards of living and therefore show increasing poverty amongst individuals (Lucifora, 2001).

Some concerns of the consequences of wage dispersions are related to individuals located at the bottom end of the earnings distribution who are more strongly affected in terms of poverty and social exclusion. Differences in wage inequality are also linked to the low wage employment problem. Countries with wide wage differentials are often characterised by a larger proportion of low paid individuals (those whose earnings fall below two-thirds of the median wage).

Hence, understanding the determinants of wage inequality is the key to understanding how to increase the standards of living of individuals in the world, and thereby to lessen poverty.

While there is a vast amount of literature investigating the reasons and patterns of wage inequality, the persistent increase of this phenomenon in certain countries over the past thirty years still attracts lots of attention. There is substantial agreement among researchers with reference to some of the basic "facts" that need to be explained. According to Autor and Katz (1999), in the U.S., changes in wage structure over the last few decades have been characterised by an expansion of wage dispersion within demographic and skill groups. Wage differentials have particularly increased by education, occupation and age (potential experience) groups; from the end of the 1970s to the mid-1990s, wage dispersion for both men and women experienced a substantial increase.

Empirical evidence (Machin and Van Reenen, 2008) documents that most changes in wage inequality in both the UK and the US have taken place from 1979. To illustrate the magnitude of changes in both the US and the UK wage structure, figure 1.1 and figure 1.2 plot respectively for the US and the UK, the evolution of the 90-50 log earnings ratio and 50-10 log earnings ratio since the 1970 until the late 2000s.

Figure 1.1 (taken from Goldin and Katz, 2007) documens the divergence in inequality trends in the US between the top and the bottom half of the wage distribution between 1974 and 2005. The figure contrasts trends in the 90-50 and 50-10 log hourly wage differentials for all workers showing the substantial increases in wage inequality occurred in both the upper half (90-50) and the

lower half (50-10) of the distribution from 1979 to 1987. After 1987 the trends in the upper and lower half wage inequality diverged, with the former continuing to rise steadlilt and the later ceasing to rise, contracting by 4 log points from 1987 to 2005.

Figure 1.2 (taken from Machin, 2010) plot the evolution of the 90-50 log earnings ratio and 50-10 log earnings ratio since the late 1970 until the late 2000s for full-time male workers in the UK. As documented by the figure, the upper tail wage inequality (90-50 wage ratio) rises sharply from the late 1970s, and consistently throughout the entire period up until 2009. The 50-10 wage gap also shows a significant increase, but most of its increase is concentrated in the 1980s and the early to mid 1990s. Following that it flattens out.

Comparing the experiences for both countries, the 1980s shows a clear picture: wage growth was more pronounced at higher points of the distribution and this is almost monotonic in both countries, leading to large increases in wage inequality. The picture is more complex post 1990. In both countries the 90-50 continues to diverge whereas the 50-10 in the US actually shrinks, indicating wage compression. In the UK the 50-10 is stable, increasing a bit in the 1990s and shrinking a bit in the 2000s. Overall then, the increase in wage inequality has been stronger in the upper tail than the lower tail taking the period as a whole and has been more pronounced in the 1980s than post 1990.

The substantial existing literature on wage inequality does not reach any consensus on what has caused this secular rise in wage dispersion across countries and in particular what has been driving the divergence between the upper and lower tails of the wage distribution. Researchers agree that the causes seem to lie with a variety of components, rather than one exclusive factor, that jointly affect the wage structure.

The main explanations proposed to describe the rise in wage inequality include: demand-side explanations; the supply-side explanations; changes in the composition of skills, the role of trades and the Skill Biased Technological Change. In research that tries to reconcile cross-country differences in change in wage inequality, an emphasis has been placed upon the role of labour market institutions that affects wage structure differently in different countries.

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, 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 "…is generally believed to account for most of the growth in overall wage inequality" (Lemieux 2006 page 461). As pointed out by Lemieux (2006), the standard human capital variables like education and experience explain only about one third of the variance of wages. Intuitively, years of schooling and experience do not capture returns to other skills. By contrast, regression-based residuals include unmeasured aspects of human capital such as school quality, ability, effort or innate skills.

One factor that has been recently offered as a potential explanation of increasing residual wage inequality is the changing age and educational composition of the labour force. However, the existing literature investigating the effect of changes in the education and experience composition of the labour force on the evolution of the residual (JMP 1993, AKK 2005, 2008, Lemieux, 2006) in part because of different methodologies, provides (different) results that are relatively difficult to compare.

The first chapter of my empirical analysis (Chapter 2) is concerned with reassessing the evolution over time of the two wage components (between and within-group) in a comparative framework for Italy, the UK and the U.S. by applying the same methodology. Specifically, Chapter 2 seeks to test the validity of Lemieux's composition effect hypothesis (2006), i.e. that composition effects exert an upward mechanical force on the residual wage inequality. Therefore, Chapter 2 aims to evaluate whether changes in the education and experience of the work force that occurred during the last few decades, can help to account for the increase in residual wage dispersion in the UK and Italy, as appears to be the case in the U.S.

This comparison is motivated mainly by the fact that the majority of evidence documenting the role of the residual in the overall wage evolution is based on the U.S. experience. Less evidence exists for both Italy (Naticchioni et

al., 2008) and the UK (Gosling et al. 2000, Machin and Van Reenen 2008). Moreover, the divergent experience of Italy, the UK and the U.S. in terms of wage institutional pay settings and inequality makes the comparison particularly fruitful. Italy differs from both the UK and the U.S. in the more centralized wage structure; in Italy collective bargaining agreements, generally at the industry level, set minimum rates for the lowest pay group in a collective agreement. The three countries have also been characterised by the labour force growing older and more educated, though at different rates across these countries.

Over the last few decades, the UK and the U.S. have not only experienced notable increases in the degree of wage inequality, but in both countries immigration has increased significantly.

There is a huge empirical debate on the social and economic consequences of international migration, one of the core concerns related to the impact of immigration on the wages of native workers. Despite the common-sense intuition behind the theoretical implications of the laws of supply and demand, the international migration literature has struggled to arrive at a consensus on the impact of immigration on the wages of workers in the receiving countries.

A consensus emerging from a recent stream of the empirical literature is that immigration not only has little impact on natives' wages, but also, on average, exerts a positive rather than negative effect on natives' wages (Ottaviano and Peri 2006; Manacorda, Manning and Wadsworth 2007; Dustmann, Frattini, Preston 2008).

The aim of Chapter 3 is to adapt one of the main challenges of the 1990s wage literature to the immigration context: wage dispersion is not fully explained by variables linked to the standard human capital model, like education and experience. Residual or within-group wage inequality – wage dispersion among workers with the same education and experience - accounts for most of the growth in overall wage inequality (Juhn et al. 1993; Acemoglu 2002; Autor et al. 2005, 2008; Lemieux 2006). Therefore unlike previous studies, Chapter 3 of this thesis focuses on the effects of immigration on residual wage inequality in the UK and the US between 1994 and 2008. It seeks to assess whether and to what degree

immigration contributed, along with technology, institutions and traditional explanations, to widening inequality.

Wage inequality is often measured as the difference between the wage of workers with a college degree and those with a high school degree. In a recent contribution, Moretti (2010) questioned the relative real wage increases for U.S. 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 areas.

Chapter 4 aims 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.

The returns to graduate education in the UK has increased since the 1970s; while were high and either increasing or stable in the 1980s and the 1990s (Machin, 2003).

Figure 1.3 (taken for Green and Zhu, 2010) plots quantile regression estimates of the returns to graduate education in Britain, using consistent male employees data taken from the UK Quarterly Labou Force Survey (QLFS), 5th wave only. The estimates give the log pay increase associated with the difference between achieving GCSE grades A_C or equaivalent (Level 2) and graduating from tertiary education with at least a college degree or professional qualification (level 4 or above). The figure presents a "before and after" picture of the effects of the participation surge on the dispersion of returns to graduates quantiles rose from 0.01 to 0.11 log points for men.

The proportion of graduates in the UK labour force has risen from 9% to more than 13% over the last 15 years to 2006 (Walker and Zhu, 2008). By the beginning of the current decade several years of the now-larger graduate cohorts had entered the labour force, replacing retiring cohorts with much lower levels of educational achievement (Green and Zhu, 2010). Figure 1.4 (taken by Green and Zhu, 2010) plot the share of graduates in the labour force, considering the proportion of 25-60 year-old workforce with qualifications at level 4 or above between 1994 and 2007. The figure documents that the proportion of graduates in

the labour force grew rapidly, the supply of female graduates in particular appearing to accelerate after 2002.

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 Price Index (RPI), 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 British regions. As such, differences in regional housing costs might be expected to play an important part in determining cost-of-living differences between regions. This implies that deflating the nominal wage by a regional RPI might lead to different estimates of the observed real wage dispersion; on the other hand, one based on a national index might fail in being fully representative at the regional level.

In Chapter 4, I 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 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.

This thesis therefore provides a thorough consideration of the role of compositional changes of the labour force, the increase of immigration and the regional differences in the cost-of-living on the increasing wage inequality in the UK and in other countries from the late 1980s to 2008.

In next chapter, I begin by investigating to what extent changing characteristics of the labour force can help to explain the fact that residual or within-group wage inequality is generally believed to account for most of the growth in wage inequality. I compare the results for men and women in the UK, the US and Italy from 1987 to 2003.

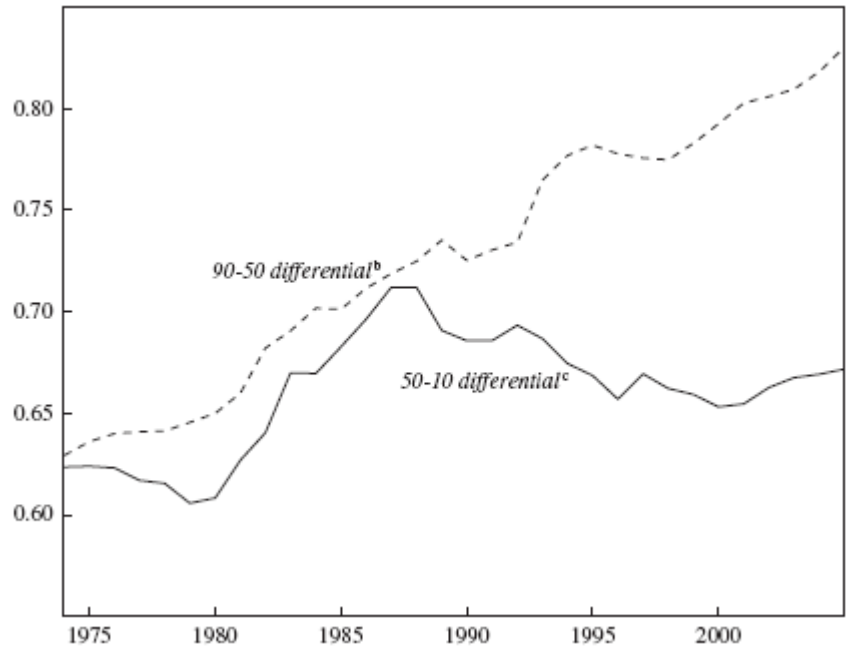
Chapter 3 of the thesis focuses on the effects of immigration on the residual wage inequality in the UK and US between 1994 and 2008. It seeks to

assess whether and to what degree immigration contributed, along with technology, institutions and traditional explanations, to widening inequality.

In Chapter 4 of this thesis, my aim is to assess how existing estimates of inequality for the UK change when differences in the cost-of-living across the British regions are taken into account.

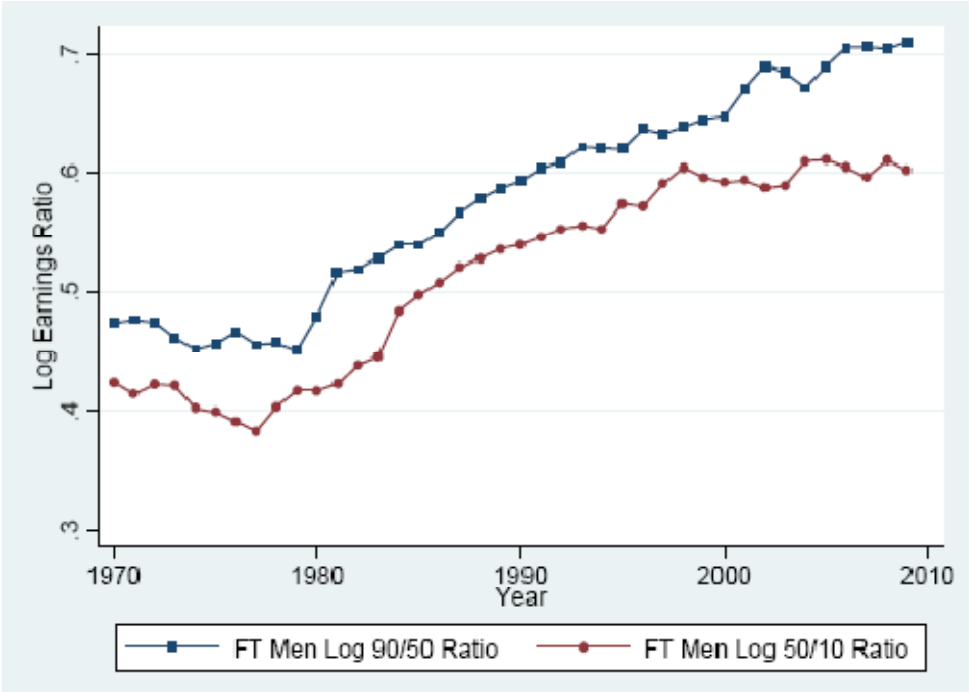
The conclusions of chapter 5 enable me to address future research.

Figure 1.1 Changes in the US wage structure, 1974-2006.



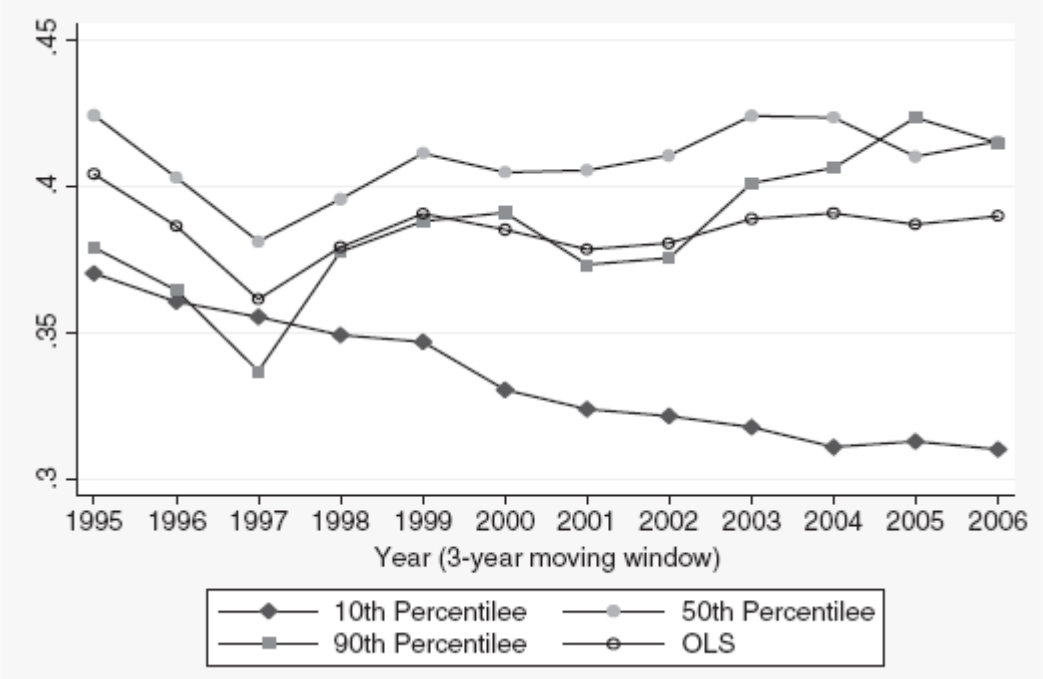
Source: Goldin and Katz, 2007.

Figure 1.2 Changes in the UK wage structure, 1970-2010.



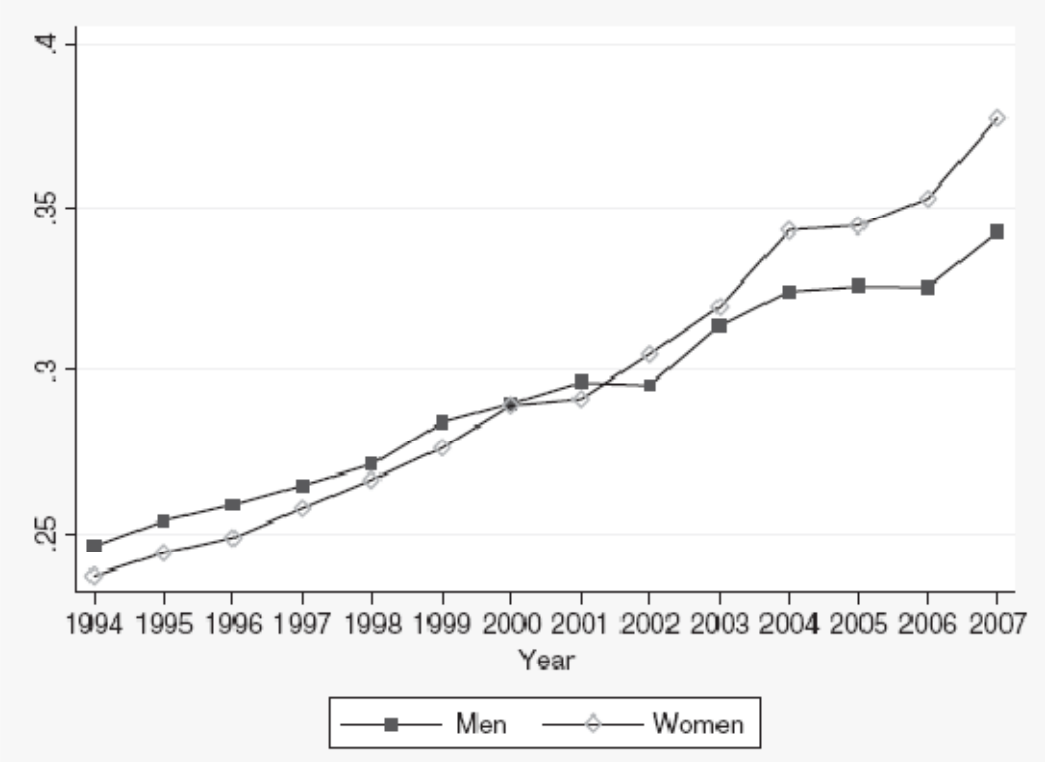
Source: Machin, 2010.

Figure 1.3: The evolution of the returns to education in the UK, 1995-2006.



Source: Green and Zhu, 2010.

Figure 1.4: The expansion of graduates in the UK, 1994-2007.



Source: Green and Zuh, 2010.

Chapter 2: Increasing Wage Inequality: a Comparative Analysis between Italy, the UK and the U.S.

2.1 Introduction

While there is an abundance of literature investigating the reasons and patterns of wage inequality, the persistent increase of this phenomenon in certain countries still attracts lots of attention from economists. The majority of the literature focuses on the U.S. where wage inequality increased most, and on the UK. By contrast, wage inequality in major European countries like Germany and Italy is characterised by a more stable wage structure over time, although it has been rising slowly recently (Blau and Kahn 1996, Gosling and Lemieux 2004, Machin and Van Reenen 2008).

Wage inequality in the UK has increased sharply since the late 1970s. Between the late 1970s and the 1990s the extent and pattern of wage inequality in the UK became increasingly similar to that in the U.S. (Gosling and Lemieux, 2004). Wage differentials in Italy, like some other European countries (e.g. Germany and France), show an overall u-shaped evolution which contrasts with the steady increase of inequality in the UK and the U.S. Wage dispersion in Italy fell between the late 1970s and the mid-1980s and rose thereafter. The dispersion of log hourly wages remained moderately flat between 1993 and 2002, and decreased during the last years of the sample for men, whilst it decreased between 1993 and 2000 for women.

The existing literature on wage inequality does not reach any consensus on what caused this secular rise in wage dispersion across countries and in particular what has been driving the divergence between the upper and lower tails of the wage distribution. Researchers 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, "…is generally believed to account for most of the growth in overall wage inequality" (Lemieux 2006 page 461). As pointed out by Lemieux (2006), the standard human capital variables like education and experience only explain about one third of the variance of wages. Intuitively, years of schooling and experience, do not capture returns to other skills. By contrast, regression-based residuals include unmeasured aspects of human capital such as school quality, ability, effort or innate skills. An earlier study by Blau and Kahn (1996) claims that wage inequality as measured by overall wage variation or by the wage gap between workers at different parts of the distribution is affected by the distribution of skills, both measured and unmeasured and by the prices determined for those skills in the labour market.

The pioneering study of Juhn, Murphy and Pierce (1993) (JMP hereafter), documents that two-thirds of the overall increase in U.S. wage inequality can be attributed to the residual with the bulk of these increases occurring between 1970 and 1988. Similarly, Acemoglu (2002), Lemieux (2006) and Autor, Katz and Kearney (2005, 2008) (AKK hereafter), confirm that changes in the residual do account for most of the overall increase in wage dispersion even over the longer sample period used by these authors.

A number of researchers comparing the trends of between-group and within-group effects, (that is the residual) provide contrasting results. For example, AKK (2008) analyse changes in overall male wage inequality and show that both the between group wage differentials, based on the college-high school wage premium, and the residual rose together only during 1979-1987. During the 1990s, the between-group inequality continued to rise while the male 90/10 residual inequality stabilized. JMP (1993) find a substantial difference in the timing of the increase in the between and within component of overall wages: the between component increased from 1960 through to 1970 and fell significantly over the 1970s; by contrast the within component remained stable or fell over the 1960s and then rose steadily through the end of the 1980s. The majority of evidence documenting the role of the residual in the overall wage evolution is based on the U.S. experience; less evidence exists for either Italy (Naticchioni et al. 2008) or the UK (Gosling et al. 2000, Machin and Van Reenen 2008). Naticchioni et al. (2008) found that between 1993 and 2004 within group only

plays a role in the upper tail (90-50) of the wage distribution; in particular they observed that the upper tail of the wage distribution increases, while wage compression is observed in the lower tail (50-10). Gosling et al. (2000) show that since the late 1970s the increase in within-wage dispersion represents an important aspect of rising wage inequality for the UK.

This chapter seeks to reassess the evolution over time of the two wage components (between and within-group) in three different countries by applying the same methodology over a similar period using comparable micro data.

One factor that has been recently offered as a potential explanation of increasing residual wage inequality is the changing age and educational composition of the labour force. Over the last 25 years, the education and experience of the U.S. labour force rose substantially; AKK (2005) show that in the U.S. the full-time equivalent employment share of male workers with a college degree rose from 18 to 32% between 1973 and 2003, while the employment share of workers with lower education fell from 62 to 41%. Gains in potential experience were similarly pronounced.

The existing literature that analyses the effects of changes in the education and experience composition of the labour force on the evolution of the residual, gives relatively mixed results. JMP claim that changes in observable characteristics (education and experience) for the labour force appear "relatively unimportant"; AKK attribute only a "secondary role" to those changes in the evolution of the residual. Lemieux (2006) argues that a large fraction of the 1973-2003 growth in the residual wage inequality is a "spurious" consequence of the composition effects, in other words he shows that secular changes in the education and age structure may mechanically increase the residual wage inequality.

The second aim of this study is to test the validity of the composition effect hypothesis in countries other than the U.S. by applying the same methodology developed by Lemieux (2006) that requires taking the actual residual variance of the log hourly wage OLS regressions and re-weighting it holding the characteristics of the labour force constant at a base year. By using this technique to examine data for the UK and Italy, this chapter seeks to evaluate whether changes in the education and experience of the work force that occurred during the last decades are equally responsible for the increase in wage dispersion

in the UK and Italy, and whether they can help account for the increase in residual, as appears to be the case in the U.S.

This raises the question as to how educational and experience characteristics of the labour force evolved through time. Table 2.2 shows that there are differences in the evolution and changes of the educational level of the work force across countries. In 2003, the U.S. had the smallest share of male workers with a low level of education[1], around 10 %, while in the same year General Household Survey (GHS) data for the UK and Survey of Household's Income and Wealth (SHIW) in 2004 for Italy provides evidence that the share of male workers who only attended compulsory schooling is respectively around 56 and 45 %. The same data also show that in 2003 the share of male workers with a high level of education is greatest in the U.S., around 30%, followed by 22% in Britain and around 11 % in Italy.

As explained in part 2.2, these cross-country differences in the educational and experience levels of workers suggest that we should expect composition effects to affect with different magnitude the changes in wage inequality, however those different effects could be mitigated or exacerbated by the characteristics of the labour market institutions.

The existing literature (Blau and Kahn 1996; Lucifora 2001; Gosling and Lemieux, 2004) documents how different institutional pay settings may alter wage dispersion through various channels. Several studies show that falling unionization contributed to the steep increase in wage inequality in both the UK and the U.S. that occurred in the 1980s (Card, Lemieux, and Riddell, 2003).

Generally centralised wage-settings and institutional constraints might have a significant role in shaping the distribution of earnings across countries. Heavily regulated labour markets and highly centralized wage setting mechanisms are characterised by more rigid wage structures and lower levels of wage inequality, while less centralised wage institutions also experience the higher levels of wage inequality. This is the case for the UK and the U.S., the most

[1] Part 2.5.2 explains in detail the definition of education group used in this work. The category "low level of education" corresponds to the dropouts i.e. the workers who only completed the compulsory years of schooling; because of differences in the educational systems, especially between Italy and the U.S./UK, "low level of education" for both the U.S. and the UK corresponds to 11 years of schooling, while for the same category in Italy corresponds to 8 years of compulsory schooling (elementary and middle).

deregulated and decentralized labour markets, that historically experienced the largest increases in inequalities.

Comparing male wage inequality in the U.S. to nine other OECD countries, primarily in the mid 1980s, providing evidence that overall wage inequality is higher in the US than elsewhere, Blau and Kahn (1996) pointed out the role of institutions in affecting wage inequality. Their main finding suggests that the centralised wage-setting institutions are an important determinant of international differences in wage distributions; particularly the centralised systems of collective bargaining in other industrialised nations that increase the relative wages of workers at the bottom, i.e. reduce wage inequality compared to the U.S.

Card et al. (1999) have argued that the existence of the minimum wage, union settings, and more regulated labour markets, particularly in Western Europe, not only prevented a rise in wage inequality but also severely limited job growth. In a similar vein, Dell'Aringa and Lucifora (2001) show that wage inequality appears to be lower in countries where employment protection is stricter. The Italian experience can be quite informative on this point: Italy is characterised by a strong employment regulation and a rigid system of wage determination[2] and exhibits one of the highest degrees of wage compression in Europe. Manacorda (2004) proves that the abolition of the wage indexation system (*scale mobile)* in 1993 were one of the main factors determining the relatively recent u-shaped evolution of earnings inequality in Italy.

The divergent experience of Italy, the UK and the U.S. in terms of wage institutional pay settings[3] and inequality makes the comparison particularly fruitful. Italy differs from both the UK and the U.S. in the more centralized wage structure. In Italy collective bargaining agreements, generally at the industry level, set minimum rates for the lowest pay group in a collective agreement. Moreover, in Italy, master industry wide agreements negotiated at the national level between unions and employer associations have traditionally been the norm. In the late 1980s in Italy, bargaining shifted in some cases from the industry to the plant level.

[2] For example the "Charter of Workers" (Statuto dei Lavoratori) established various employment restrictions for firms with more than 15 employees on individual hiring and firing procedures as well as on temporary labour contracts (Lucifora, 2001).

[3] See Blau and Kahn (1996) for an extensive overview of international differences in labour market institutions.

Wage setting is more decentralised in both union and non union sectors in the U.S., though deunionisation proceeded quite rapidly. Collective bargaining in Britain is less centralised than in other countries but more centralised than in the U.S. The UK, like the U.S, had a mixture of single-firm and multiemployer agreements in the 1970s and bargaining appeared to be more centralised in the UK. For example in 1977-78, 25% of manufactured contract covered more than one firm, whereas for the U.S. the figure in 1975 was only 13%. Both countries remain less centralised institutions while the Italian system of collective bargaining remains more centralised.

The analysis of wage inequality in Italy also complements the existing literature by providing insights into a country characterised by a more regulated labour market and a more rigid wage structure. Italy could provide an image of whether and how much changes in the composition of the labour force contribute to variations in residual inequality in countries in which the level of wage inequality is lower, albeit rising slowly over the sample period. This chapter also contributes to the existing Italian literature by analysing the effect of compositional changes in variations in within-group inequality in a cross-country analysis, going back to the earliest year possible and by applying a different methodology to the relevant contributions.

This chapter adheres to the methodology and data design used by Lemieux (2006), focusing on the cross-country trends and timing of residual wage inequality to investigate if and how much changes in the composition of the labour force contributed to the increasing dispersion of residual wage inequality in countries characterised by different institutional settings. This research illustrates and analyses the evolution of the upper and lower tails of the distribution, controlling for composition effects and evaluating the price of unobservable skills. The previous literature analysing residual wage inequality (JMP 1993, AKK 2005, 2008) in part because of different methodologies, provides (different) results that are relatively difficult to compare. The methodology used by Lemieux that is implemented in this study is based on a simple approach (variance decomposition) that facilitates decomposition of changes in the distribution of wages over time into the different components that can be economically interpretable using the standard tools of the human capital model. In addition, the technique controls for changes over time in the distributions of observed characteristics of workers by

computing empirically a counterfactual distribution that emphasises the effect of compositional changes in within-group inequality.

The chapter is organised as follows: part two provides the main empirical evidence related to the residual wage inequality and composition effects; part three recalls the theoretical model behind the research; part four describes the three different datasets; part five explains the econometric methodology and the re-weighting approach originally applied by Lemieux; part six presents and compares the empirical results for Italy, the UK and the U.S.; and part seven concludes.

2.2 Wage Inequality, Residual and Composition Effects: Evidence and Issues

Between the end of the 1960s and the end of the 1980s, the variance of log real weekly wages for full time male workers in the U.S. increased by about 72% (JMP 1993). Over the last three decades, the secular increase in overall wage inequality observed mainly in the U.S. (but also in the UK) largely demonstrated two main patterns: a persistent rising wage dispersion within narrowly defined education-experience groups and a divergence in the trends of upper and lower path inequality.

Evidence from the U.S. (Autor and Katz, 1999; Katz and Murphy, 1992; Levy and Murnane,1992) shows that residual wage inequality started increasing in the 1970s and continued to rise considerably in the 1980s, reaching a level in 1987 that was 30% greater than the 1970 level, and then rose at a slower pace in the 1990s; the residual log weekly wage inequality for full time, full year workers increased by 27 log points for men and 25 log points for women from 1963 to 1995. More recently, Lemieux (2006) outlines the main patterns observed in the United States using Current Population Survey (CPS) data: the residual variance for full time, full year male workers from 1973 to 2003 grew by about 0.04 log points; most of that growth was concentrated in the 1980s. Levels remained essentially unchanged during the 1990s but grew again between 1999 and 2003.

Important findings also suggest that the increase in wage dispersion has been higher for the highest paid workers, while workers at the bottom of the wage distribution experienced less variation. Katz and Murphy (1992) show that in the U.S. the log wage gap between the ninetieth and tenth percentile within

experience-education groups increased by approximately 0.26 for men and 0.21 for women from 1963 to 1987. The change in the standard deviation and the 90-10 log wage gap is three times larger between the 1970s and 1980s than between the 1990s and 2000s (Lemieux 2006). The earlier influential work of JMP (1993) provides evidence that the diverging pattern between less and more skilled male workers in the U.S. goes back to the beginning of the 1960s. AKK (2005 and 2008) support these findings and re-evaluate the traditional explanation for changes in the U.S. wage inequality by documenting that the growth of overall wage inequality during the 1990s hides divergent patterns of inequality between the upper and lower segments. There has been a persistent rise in upper tail inequality and a decline of inequality in the lower tail since the second half of the 1980s[4]. By using MARCH CPS data from 1963 to 2005, they show that the 90th percentile rose by approximately 45 log points relative to the 10th percentile for both men and women.

They show that upper and lower tail wage inequality expanded rapidly in the first half of the 1980s for both men and women[5]. The 50-10 wage gap for the most part stopped growing after 1987. Male 50/10 residual wage inequality rose by 5.7 log points between 1973 and 1989, and then fell by 1.3 log points between 1989 and 2005. By contrast, the U.S. 90-50 gap for both men and women in 2003 was 40% larger than in 1973. The 90-50 wage gap continued to grow smoothly from 1979 to 2005; from 1973 to 1989 the male 90-50 residual wage inequality rose by 4.4 log points and by 4.0 log points from 1989 to 2005. Machin and Van Reenen (2008) explain that the increase in the 90-50 wage differential since the late 1970s occurs not only in the U.S. but also in the UK; whereas in the UK the 50-10 is broadly stable, increasing a little in the 1990s and shrinking a little in the 2000s. According to Gosling and Lemieux (2004), UK wage evolution between 1979 and 1998 was almost entirely driven by changes in within-group inequality which increased more for men (0.073) than women (0.056). Similar to the US, the bulk of this increase occurred during the 1980s. Gosling et al. (2000) demonstrate

[4]They illustrate that a falling minimum wage contributed to raising lower tail (50-10) wage inequality in the 1980s.

[5] In a recent contribution, Dustmann, Ludsteck and Schonberg (2009) revisited the changes in the wage structure in West Germany and demonstrated that wage inequality has increased in the 80s, but only at the bottom part of the distribution; they also find that it is important to account for changes in education and experience of the labour force and this is true especially for the top half of the wage distribution.

that like the US, since the 1970s an important aspect of rising wage inequality in the UK is increased within-group wage dispersion.

The recent evidence for Italy shows that residual wage inequality has been increasing over time. For example, Lilla and Staffolani[6] (2007) demonstrate that within-group inequality increased for daily wages from 1985 to 1999. Devicienti (2003), using administrative data from 1985 to 1996, examines how much of the increase in earnings inequality can be attributed to changes in individual characteristics, the price attached by the labour market, and the effect of unobservables. He finds that inequality is mainly explained by its within-group component. Lilla (2005) carried out a within-between analysis of wage inequality in Italy, using the SHIW data he claims that both between and within components slightly increased between 1998 and 2002. By contrast, Naticchioni, Ricci and Rustichella (2008) using SHIW data from 1993 to 2004 found that within group only plays a role in the upper tail of the wage distribution; in particular they observed that the upper tail (90-50) of the wage distribution increases, while wage compression is observed in the lower tail (50-10).

This paper contributes to the existing Italian literature by controlling for characteristics of the labour force and by comparing them to the more striking experience of the U.S. and the UK with regard to rising wage inequality.

Several researchers document that much of the increase in wage inequality since the 1970s is due to a dramatic increase in the demand for more educated workers and "more skilled" workers (Katz and Murphy 1992). This confirms the prediction of human capital models that wage inequality can increase because the returns to education and experience increase or because the demand for unobserved skills increases. Becker (1964) and Card (2001) argue that there is no single "return to education" but rather distributions of returns across heterogeneous individuals. In recent work, Lemieux (2006a) confirms the findings of Mincer (1997) and Deschènes (2002) that the return to post-secondary education increased sharply while returns to lower levels of education remained unchanged. In particular, using a quantile regression approach, he shows that the return to post-secondary education has increased more in the upper quartile of the wage distribution.

[6] Lilla and Staffoloni (2007) analyse the evolution of inequality in yearly and daily wages between and within groups of blue and white collar, using the INPS-ISFOL database.

Martins and Pereira (2004), using a quantile regression, analysed the dispersion of the returns to education for 14 western countries during the mid-1990s, addressing the link between and within-levels inequality. They found that the earnings increment associated to schooling is higher for those individuals whose unobservable characteristics place them at the top of the conditional wage distribution, implying that schooling may have a positive impact upon within-group wage inequality, as the spread of the returns increases for higher educational levels.

During the 1990s, numerous researchers strongly supported the idea that Skill Biased Technological Change (SBTC) was one of the leading explanations for the increase in wage inequality. As explained by Acemoglu (2002), technological change complements the demand for more skilled workers, and substitutes tasks performed by the unskilled, thus raising inequality. Katz and Autor (1999) propose a similar conclusion, demonstrating that SBTC was one of the driving forces behind the increase in the relative demand for skills[7]. They argue that the utilization of more skilled workers is positively correlated with capital intensity and the implementation of new technologies, both across industries and across plants within detailed industries. These patterns indicate that physical capital and new technologies complement more skilled workers; therefore secular increases in the capital/labour ratio can be considered as a source of secular growth in the relative demand for skilled labour. Another indicator of the SBTC as a demand for skill driving force can be seen in the strong correlation between industry-level indicators of technological change (computer investments, the growth of employee computer use, research and development (R&D) expenditures, utilization of scientists and engineers) and the within-industry growth in the relative employment and labour cost share of more skilled workers. Bound and Johnson (1992) and Katz and Murphy (1992) argue that because of SBTC, there has been an increase in the demand for more educated workers exceeding the relative supply and causing the wage of college educated workers to increase. Machin and Van Reneen (2008) claim that SBTC, together with institutional changes, better explain the increase in the upper tail of the

[7] Katz and Autor also explain how computers may influence relative labour demand in several ways. Computer business system often involves the routinization of many white collar tasks. Simple repetitive tasks have proved more amenable to computerization than more complex tasks. Direct substitution of computers for human judgement and labour is likely to have been more important in clerical and production jobs than in managerial and professional jobs.

distribution for the UK and the U.S. since the late 1970s. In the 1980s the 90-10 percentile tail expanded by 1.9 percentage points a year in both countries. The opposite conclusion related to the SBTC hypothesis is reached by Card and DiNardo (2002) who document that the timing of the growth in the wage inequality is difficult to reconcile with the SBTC hypothesis because the rate of technological innovation was stronger in the 90s when wage inequality grew less.

These leading existing explanations on wage inequality such as declining unionization, the falling real minimum wage and the SBTC[8] do not then seem to help explain the main recent trends in wage inequality, since they find it harder to predict increasing upper-tail inequality and compressing lower tail inequality (AKK,2005). These explanations also fail to explain the fact that "residual […] is generally believed to account for most of the growth in overall wage inequality" (Lemieux 2006 page 461). As documented by Lemieux, there are two main reasons affecting the increase in residual wage inequality over the last decade:

i) The "price" or return to unobserved skills may be increasing over time because of the increase in the demand for skills;

ii) Dispersion could be increasing because of composition effects.

In an earlier contribution, Levy and Murnane (1992) present a set of hypotheses for explaining not only within-group inequality but also the growth of within-group variation over time. Their hypotheses include both supply and demand shifts for workers characteristics; the former consists of the changing characteristics of the labour force (including aptitude test scores, measures of ability to work with other people); as well as increasing returns to skill; the latter includes plant-specific wage differentials within industry as well as changes in wage-setting institutions.

Green and Zhu (2010) provide evidence that dispersion of the returns to graduate education in the UK substantially increased for both men and women over the 1994 to 2006. They argue that the rising dispersion in the returns to

[8] Similarly, Lemieux (2008) discusses why the SBTC explanation presents several limitations; one reason is that it could not help to explain some diverging pattern of inequality across advanced countries; the SBTC was dependent on the weak fact that the residual inequality was increasing since the 1970s. The failure of the SBTC for other developed countries can be explained by the fact that because those countries are subject to the same technological change, they did not experience an increase in inequality, as we should instead expect. They did not vary in a similar way over time.

graduate education is associated with the increasing incidence and cost of overqualification of graduates, that is normally found to be associated with a notable pay penalty relative to those who are matched to a job at their own level.
If the rising dispersion that appeared in the recent years is associated with rising overqualification, then as long as graduates are being matched with graduate jobs, one should expect to find that the difference between the estimated effects of graduate education at the top and bottom end of the residual pay distribution remains constant. Green and Zhu (2019 also explains that one plausible implication of rising overqualification and increasing dispersion in the returns is that current and future cohorts of school leavers may perceive an increased risk of investing in higher education. Risk is also derived from uncertain course completion rates, but it is likely that the increased earnings dispersion will raise the perceived financial risks. If so, some downward pressure on enrolments would be expected.

Changes in characteristics[9] affect both the demand and supply of observed and unobserved skills and can alter wage and employment outcomes (Autor and Katz, 1999). Movements in within-group inequality may reflect market forces changing the returns to (unmeasured) skills. Therefore, the rise in within group inequality can be interpreted as reflecting a rise in the returns to unobserved skills. Holding market prices constant and changes in labour force composition can mechanically raise or lower overall earnings dispersion by increasing or reducing heterogeneity in observed skills (education and experience).

An increase in the proportion of the workers with more education and experience can mechanically raise residual wage inequality also because earnings variation is higher for those with college education relative to high school education (Machin and Van Reenen, 2008).

The link between composition effects and the residual can be explained by the fact that when the level of education of the labour force increases, there are more and more “marginal“ workers added to the high-education workers group, creating more unobserved heterogeneity in that group and increasing within-group inequality. Education and unobserved skills are imperfectly related: there are skilled and less skilled workers within the same education group. Therefore, an

[9] Acemoglu (2002) explains that composition effects cannot by themselves explain the recent changes in wage dispersion; but it suggests that inequality amongst more educated and less educated workers should move in opposite directions.

increase in the supply of more educated workers will immediately benefit workers with more unobserved skills, and will also depress returns to schooling whilst raising within-group inequality. In a similar vein, Martins and Pereira (2004) explain that one reason behind the positive effect of education on within-group inequality is the complementarities between schooling and unobserved ability, in which the most able can benefit more from their schooling; in general, education interacts with a set of factors that are heterogeneously distributed across workers within the same education level. Based on the same line of argument Green and Zhu (2010) point out that graduate education is complementary with unobserved ability, implying that those graduates at the lower end of the unobserved ability distribution are less able than their counterparts in earlier cohort, and hence gained fewer benefits from graduate education. In similar vein Brunello et al. (2009) confirm that unobserved ability and the labour market luck are substitutes to education in the production of human capital and earnings.

In a single index model, in which there is only one type of skill which is imperfectly approximated by education/experience, observed and unobserved skills are imperfect substitutes. An increase in the returns to observed skills (such as education) will also be associated with an increase in the returns to unobserved skills. This means that between and within-group inequality should move together (Acemoglu, 2002). In a two-index model where observed and unobserved skills are imperfect substitutes (Acemoglu 2002), skills are multidimensional; assuming that more skilled workers within each education group also benefit from skill-biased technical progress, technical change spurred by the increase in the supply of educated workers will immediately benefit workers with more unobserved skills, raising within-group inequality. Therefore, an increase in the supply of educated workers will depress the returns to schooling, while increasing within-group inequality[10].

Evidence of this is given in tables 2.1 and 3.3a-3.3f which report changes in the share of workers by education/experience and the corresponding changes in within-group inequality. This confirms, only in part, one of the main findings of Lemieux. The results show that an increase in the supply of more educated

[10] Mincer (1997) applies a human capital analysis to intra-group wage inequality, measured by variances in log-wages, and their changes over time reveal to the U.S. wage structure changes from 1970-1990. In a similar vein to Acemoglu, he provides evidence that within-group inequality is not directly or closely related to between group variances, therefore we can expect difference in movements in the two components of inequality.

workers does not always generate an increase in within-group inequality and a decrease of the supply of less educated workers does not always correspond to a decrease in within-group dispersion. This appears to be particularly evident for Italy.

A number of recent influential studies focus on this interplay between changes in educational and experience characteristics of the labour force and the evolution of the residual to evaluate how much of the overall increase in wage inequality can be attributed to wage dispersion among workers with the same education and experience; and how much of the increase in residual wage inequality is due to changes in the composition of the work force.

Although the main contributions in this (AKK 2005 and 2008; JMP, 1993 and Lemieux 2006) use the same date source (Current Population Survey (CPS), United States) the conclusions they offer are relatively mixed.[11]

JMP (1993) use CPS weekly and hourly wages for males from 1963 to 1989 to analyse how much of the rise in residual wage inequality is due to related increases in the market return to skill, given that the rise in skill premia applies to both observable (education, experience and occupation) and unobservable dimensions of skill (the residual). They show that unobserved components affect with different magnitude the wages of workers at the top and at the bottom of the distribution. Over the period 1964-1988, changes in unobservable quantities accounted for 65 % of the increase in inequality for workers below the median but less than half of the increase in inequality for those above the median. JMP analyse the composition effects on the level of wage dispersion, concluding that changes in observable characteristics of the labour force appear "relatively unimportant"[12]. Another striking finding of JMP relates to the timing of the increase between and within groups. The authors prove that not only the rise in inequality and the rise in education premia are distinct phenomena, but they are also characterized by a difference in timing of the increase in wage inequality within and between groups. From 1960 to 1970 wage differentials by education increased and then fell over the 1970s. In contrast, inequality within education and

[11] In a recent paper, Lemieux (2008) highlights that difference in results might be due to differences in data; which means they are not only due to measurement errors.

[12] Acemoglu (2002) documents that on one side composition effects could provide an explanation to the recent pattern in wage inequality; on the other side he claims that the returns to education and residual inequality are not simply due to composition effect.

experience categories remained stable or fell over the 1960s and then rose steadily through the end of the 1980s.

Using the techniques outlined in JMP, Lemieux (2006) uses CPS data from 1973 to 2003, on the hourly wages of men and women, to evaluate to what extent the increase in wage dispersion in the U.S. can be attributed to the increase in residual wage inequality and how much of the increase is related to changing labour force characteristics.

Lemieux proves that the increase in residual wage inequality does not only depend on the rise in the demand for skill; but that the increase in the returns to unobserved skills account for no more than 25% of the total wage change between 1973 and 2003.

He concludes that a large fraction of the 1973-2003 growth in residual wage inequality in the United States is due to composition effects, affecting both the upper and lower tails of the distribution. In particular he demonstrates that the increase in within wage inequality appears to be a *spurious* consequence of the fact that workforces became older and more educated over time; in other words the increases in educational level and experience act as a mechanical force on the residual such that the increase in the variance of the residual follows the same sign of the change in composition of the work force. This means that changes in education and experience characteristics of the work force determine more variation in wages due to unmeasured aspects of human capital.

AKK (2005) use a quantile decomposition technique proposed by Machado and Mata (2005) to reconsider Lemieux's composition hypothesis, i.e. the role of changing labour force composition (education and experience) and changing labour market prices to the expansion and divergence of upper and lower tail inequality. This methodology partitions the observed distribution of earnings into "price" components (wage coefficient) and "quantity" components (labour force composition) and calculates, through simulation, the impact of each change in overall wage dispersion. In doing so, AKK simulate the counterfactual distribution of wages that would have prevailed if labour force composition were as in time period t and labour market prices were as in time period r. This simulation captures the effects of composition on both between-group and residual inequality, analysing the contribution to changes in wage inequality both above and below the median of the distribution. They show that even though

changes in labour force composition exerted an *upward force on* residual wage dispersion, *they* do *not contribute by themselves* to an explanation for the diverging path of upper and lower tail inequality. In the U.S., composition changes of the labour force only affect the lower tail wage distribution; upper tail inequality grows steadily, accounting for over three-quarters of total male inequality growth from 1975 to 2003.

AKK show that the 90/50 residual wage gap inequality rose by 8.4 log points from 1973 to 2005. Holding labour force composition constant at its 1973, 1989 or 2005 levels does not change the basic message: the composition constant rise in residual 90/50 inequality is at least 65% as large as the actual residual. In a similar vein to JMP, they demonstrate that education and experience characteristics of workers affect with different magnitude skilled and unskilled workers. The steady growth of upper-tail inequality during the last two decades is not due to mechanical effects of composition, but to changing labour market prices. By contrast, composition can fully explain the overall trend in residual inequality during the 1990s but only plays a secondary role when considering the upper and lower tails separately.

The methodology applied by AKK (2005) allows deriving counterfactual wage distributions such that the changes over time of the wage distribution are decomposed into price, quantities and residual. Similarly, the methodology applied by Lemieux and that this paper uses, not only allows a decomposition of the total change of wages over a period into price, quantities and residual but also accounts for changes in observed characteristics by emphasising the effect of compositional changes on variations in within-group inequality.

In fact, the key insight of Lemieux (2006) is that the approach controls for changing education and age characteristics of the labour force by holding them constant at a base year in order to untangle their effects on the evolution of the residual[13] and therefore on the overall wage dispersion. The technique, that requires taking actual residual variance and re-weighting it according to the observed characteristics of workers in the base year, helps to detect differences in residual variance attributable to changes in observed characteristics (education and experience) in different countries.

[13] Details of this methodology are given in part 2.4.

This can be done by constructing a weight expressed as the ratio of the probability that a worker observed at the actual year is observed at a base year (conditional on the observed characteristics) to the probability that the same worker is observed at the actual year (conditional on observed characteristics).

2.3 Theoretical Framework

The human capital earnings model of Mincer (1974) provides an explanation for the dispersion in wages. Based on the model of investment in human capital, it expresses the natural logarithm of earnings as a function of years of education and years of labour market experience.

The standard wage equation takes the following form:

1) $\log w = c + rS + b_1E + b_2E^2 + \varepsilon$

Where c is a constant, S is years of schooling; E is years of potential labour market experience and ε is the standard regression residual. An additional contribution of the Mincer model is the introduction of potential labour market experience (derived as the difference between the age and years of schooling minus the age at which children generally start schooling) as a regressor in the earnings equation.

Almost all studies (Mincer, 1974, Deschenês, 2002; Levy and Murnane, 1992, Welsh 1979; Borjas 2007) show that schooling has a positive and significant effect on earnings ($r>0$) and that earnings are a concave function of labour market experience ($b_1>0$ and $b_2<0$). In his earlier work Mincer (1958) pointed out that the "age-earnings profile" was steeper for more educated workers than for less educated workers. In other words, log earnings are not a strictly separable function of education and age; meaning that there are different rates of return for each age group, rather than a single rate of return to education.

The equation (1) can be interpreted as the outcome of a process by which individuals invest in two types of human capital: education and on-the-job training (OJT). Education is directly measured by years of schooling, while years of labour market experience is a proxy for on-the-job training.

According to this model, we would expect the "age-earnings profile" to be steeper for more educated workers than for less educated workers. Individuals who invest more in education have higher marginal returns to education than others, therefore the log wage-schooling relationship would be convex, and the labour market "price" of schooling will be higher at higher levels of education.

Given equation (1), the error component model for the residuals ε can be expressed as the product of some unobserved skills *e* with associated returns *p*:

(2) $\varepsilon = pe$.

From equation (2) it is possible to measure inequality in the residual ε, by using standard measures of dispersion such as the variance or percentile differences. In a standard human capital model, wage inequality can increase either because the returns to education and experience increase, or because residual inequality increases.

2.4 Econometric Methodology and Reweighted Approach à la Lemieux

Since this work seeks to test Lemieux's findings, the methodology and the identification strategy largely adhere to his original work, although some modifications and adjustments of data sets have been necessary. To assess the contribution of observable and unobservable components of wage dispersion to changes in overall wage inequality, this work follows part of the existing literature (Autor and Katz, 1999; JMP 1993, and Lemieux 2006) which applies the standard variance decomposition.

The econometric methodology of this work is based on two simple steps: the analysis of inequality in the residuals and a reweighting approach to control for compositional changes.

The residuals are obtained from a standard OLS regression, having the following specification:

(3) $$y_{it} = X_{it}B_t + \varepsilon_{it}$$

Where y_{it} is the log hourly wage of individual *I* in year *t*, X_{it} is a vector of observed individual characteristics (education, age and a set of interaction terms

between education and age), B_t is a vector of estimated returns to observable characteristics in *t,* and ε_{it} is the log wage residual depending on unmeasurable skills.

Given the orthogonality of the predicted values and the residuals in an OLS regression, the variance of y_{it} can be written as:

$$\text{(4)} \qquad \text{Var}(y_{it}) = \text{Var}(X_{it}B_t) + \text{Var}(\varepsilon_{it}).$$

In other words, the change in the variance of log wages can be decomposed into the change in the variance in the predicted values (between-group inequality), reflecting the contribution of observable prices and quantities; and the change in the residual variance (within-group inequality) measuring the role of unobserved skills.

Changes in the residual variance can be attributed to changes in prices for unobserved skills (p_t^2) and changes in unobservable skills (var e_{it}) if $\varepsilon_{it} = p_t e_{it}$ then

$$\text{(5)} \qquad \text{var}(\varepsilon_{it}) = p_t^2 \, \text{var}(e_{it}) .$$

This dispersion can be disaggregated across time and across education and experience groups. Because the information and structure of the data sets used in this work (ORG/CPS, GHS and SHIW) are somewhat comparable, it is possible to divide workers into 12 education and experience cells based on a group of 3 (similar) education categories (lower, intermediate and high) and 4 potential experience categories 1-10, 11-20, 21-30 and 31+ years.

To control for composition effects, i.e. to assess if and how much changes in the educational level and experience of the labour force account for the increase of wage inequality, the variance of the residual is recomputed, assigning workers observed at the actual year the same characteristics of workers observed at a given base year; in other words the technique holds the skill composition of the work force (θ_{jt}) constant over time in order to reflect the distribution of characteristics of the labour force at a given year.

Assuming[14] that observed skills, x_{jt}, can be divided into a finite number of education-experience groups j, then the (unconditional) variance of the residual

[14] This part follows Lemieux (2006).

$Var(e_{it})$ is directly affected by changes in θ_{jt} (share of workers in experience-education group j at time t) and is linked to the conditional variance σ^2_{jt} by the following:

$$(6) \qquad var(e_{it}) = \Sigma\, \theta_{jt}\, \sigma^2_{jt}$$

where σ^2_{jt} represents the conditional variance ($Var\,(e_{it}|x_{it})$) so changes in the education-experience cell shares will correspond to changes in the residual variance. Since the conditional variance in wages V_{it} is linked to the conditional variance of unobserved skills by the following:

$$(7) \qquad V_{it} = p_t^2\, \sigma^2_{jt}$$

where p_t captures the returns to unobserved skills (e_{it}) it follows that σ^2_{jt} also increases as a function of experience and education. Only when the variance of unobserved skills remains constant over time, can changes in residual variance be interpreted as evidence of changing skill prices p_t.

Holding the characteristics of the labour force constant at a base year can be done by constructing a counterfactual weight, expressed as $(1-\omega_i)/\,\omega_i$, where ω_i is the predicted probability for each worker observed at time *t* to be in the base year *s*. This probability is obtained by applying a logit model conditioning on the characteristics (education and age) of the workers. The magnitude of the weight depends therefore on the characteristics of workers observed at the year *t* compared to the characteristics of workers observed at a given base year *s*. For example, the predicted probability of lower educated workers observed in a year characterized by a higher level of education of the workforce will be small, therefore the counterfactual weight $(1-\omega_i)/\,\omega_i$ will give more to lower educated workers in the later years. Similarly, the weight would give less weight to observations of higher educated individuals in the later sample years.

In this way, by holding the distribution of skills constant over time, it is possible to compute a counterfactual variance i.e. the variance of the residual that would prevail if the distribution of skills of workers remained constant at their base year value. The difference between the counterfactual variance and the actual variance shows how much the composition of the labour force accounts for the evolution of the observed residual.

Formally the residual variance needs to be written as a function of the variance of wages, V_{jt} within each skill group j:

$$(8) \qquad Var(\varepsilon_{it}) = \Sigma\theta_{jt} V_{jt}$$

Holding constant the skill distribution of workers (θ^*_{jt}), the counterfactual residual variance can be written as:

$$(9) \qquad V_t^* = \Sigma\theta^*_j V_{jt}$$

Working in this way will help to understand whether the composition effect is driving changes in the residual variance; in particular when the composition of the work force is held constant, any increase in the residual variance can be interpreted as an increase in skill prices p_t.

2.5 Data

2. 5.1 Description of the Samples

The data used in this study comes from individual-level micro datasets for the U.S., the UK and Italy, from 1987 to 2004. The U.S. analysis uses the May/ORG Current Population Surveys[15] (CPS); for the UK the empirical analysis is based on the time series of cross-section files from the General Household Survey (GHS); for Italy the data sources are from the Bank of Italy and the Survey of Household's Income and Wealth (SHIW).

Even though alternative datasets are available for both the UK and Italy[16], the choice of the GHS and SHIW, together with the primary data MAY/ORG CPS, is motivated by the fact that they contain similar information on individual wages, employment status and educational qualifications covering the period of interest. This makes the surveys, and the results, fairly consistent and comparable over time. Including Italian data in the analysis has been quite challenging, because of some significant differences from the other datasets. In order to ensure comparability, it has been necessary to impose some constraints on the other datasets, therefore limiting the analysis over time for both the UK and the U.S. The wage measure used is the real hourly wage rate. The GHS and SHIW do not

[15] I am grateful to Thomas Lemieux for providing me with the dataset for the U.S.
[16] Appendix provides further details of data issues.

contain a point-in-time hourly wage measure and the Bank of Italy only records information on hours worked per week since 1987.

The CPS is a monthly household survey conducted by the Bureau of Labour Statistics to measure labour force participation and employment. The survey provides individual data for approximately 30,000 individuals each month. Every household that enters the CPS is interviewed each month for 4 months, then ignored for 8 months, then interviewed again for 4 more months. Usual weekly hours/earning questions are asked only of households in their 4th and 8th interview. These outgoing interviews are the only ones included in the extracts. New households enter each month, so one quarter of households are in an outgoing rotation each month. MAY/ORG CPS provides point-in-time measures of usual hourly wage for 60% of the sample. The remaining non hourly wage has been calculated by Lemieux as the ratio of annual earnings to hours worked. Hourly wage ranges between $1 and $100, at 2003 Dollars value. Another issue addressed by Lemieux relates to topcoded wages in the CPS.

The CPS generally reports earnings information at two intervals: on an hourly basis, for "hourly workers"; and, on a annual basis, for all other, generally non-hourly, workers. The first difficulty facing efforts to make a consistent hourly earnings series is that, while hourly workers wages are generally topcoded at $99.99 per hour (a threshold rarely crossed), annual earnings are topcoded at much lower thresholds, with an important share of workers above the threshold in some years. In the release of data available to the public, the Census Bureau restricts the top of the non-hourly earnings distribution to $99,999 a year. This means that all earnings above that level appear in the CPS public use as $99,999 whatever their actual earnings are.

This artificial ceiling can lead to bias in the measurement of trends in inequality analysis because they censor the range of incomes that are observed. As pointed out by Burkhauser et al. (2007) this problem would be less of an issue when one is looking at inequality trends over time if the nature and extend of top coding were constant. However, CPS top codes have changed over time, leading to a potentially serious time inconsistency problem for inequality analysis. Based on the ORG CPS data used in chapter 3, between 1994 and 2008 the proportion of earners affected by the top coding ranges from 1.63 to 6.47 percent a year. Other things equal, any analysis of inequality is likely to underestimate any trend toward

increased earnings inequality that might have taken place at the top of the earnings distribution. In a single cross-section, topcoding will lower the mean and the variance of the wage data relative to the true mean and variance. Over time, with a fixed nominal topcode and wage inflation, the size of this bias will generally increase, lowering estimates of the change in the mean and variance of wages. With no change in the underlying mean and variance, for example, the measured mean and variance of real wages would generally fall. Irregular and large adjustments to the topcode create further problems by causing sudden jumps in the mean and variance of wages that are not related to actual changes in the true wage distribution (Schmitt, 2003).

Depending in the purpose of a specific analysis, researchers have addressed the topcoding issue in different ways. Some researchers just ignore the topcoding issue. One strategy is to attempt to estimate the mean above the topcode and assign this to mean to all topcoded observations. In principle, if all observations above the topcode are assigned the true mean above the topcode, the mean of the resulting distribution would equal the mean of the true distribution. As clarified by Schmitt (2003) this method doesn't generate the correct variance, but, in general would push the variance of the observed distribution closer to the true variance, reducing the downward bias both within a given cross-section and over time. Similarly to other researchers (Katz and Autor, 1999; Katz and Murphy, 1992; AKK, 2005, 2008) Lemieux adjusted for topcoding by multiplying topcoded wages by a factor of 1.4 which is believed to provide estimates of the mean and the variance that are closer to their true values.

In Italy, the survey is carried out every two years and 8,000 families are covered in each survey. Hourly wages have been constructed by dividing annual earnings by the product of weeks worked last year and weekly hours worked last week. One of the main and considerable features (limitations) for Italy is that the Bank of Italy only records net wage i.e. wages net of taxes[17] and social security contributions. Although this may limit the comparison across countries, requiring some caution when interpreting and extending the results for Italy; because tax is proportional to income, having net wages for Italy should not affect the trends but only the level of the results. In the SHIW, wage values that are between 1 and 50 Euros (at 2004 Euros value) are used.

[17] Italian payroll tax is a progressive system.

The GHS is a multi purpose continuous survey that collects information on a range of topics from people living in private households in Great Britain. Face-to-face interviews take place continuously throughout the year and a sample of approximately 13,000 addresses is selected each year from the Postcode Address File. The GHS contains information on weekly wages; therefore the hourly wage rate has been computed by dividing the weekly wage by the hours worked last week. Following the existing literature (Manacorda, Manning and Wadsworth (2007), wages less than £1 per hour and more than £100 per hour (at 2003 Pounds value) have been excluded.

Another discrepancy between the three data sources relates to their sample sizes; the MAY/ORG CPS sample size is more than double that of the UK GHS, and is almost four times bigger than the Italian one. The CPS set is composed of about 200,000 observations over sixteen years; the GHS of 80,000 and the SHIW 47,000 observations. This difference in size is important because it will be reflected in the size of the education and experience groups, and this could affect the accuracy of the results. Increases in sample size should improve the accuracy of the estimation, as well as possibly leading to smaller values of the residual variance. The samples are based on individuals who are full time employees[18] and main job workers, working the whole year with positive potential labour market experience.

2.5.2 Generating Comparable Education Groups

All results derive from separate regressions for men and women of the log hourly wage on a set of dummies for age, a quadratic in age, education and interactions between education and age.

Construction of a consistent variable recording the education level of workers across countries is complicated because of differences in schooling systems across the countries. In addition, the three surveys record information about schooling in a different manner and sometimes not consistently through time. The CPS contains information about years of schooling completed. The SHIW only displays information about the highest earned qualification. The GHS

[18] The use of full time workers only is meant to eliminate variation associated with hours per week or weeks per year, as well as maximize the comparability through time (Katz and Murphy, 1992; Lemieux and Card, 2001).

contains information on both earned qualifications and on ages at which individuals left school.

The Italian schooling system differs greatly from that of the United States and the United Kingdom. Children start school at the age of 6, complete compulsory school after 8 years of study, high school after 13 years and get a university degree after 4 more years[19]. In contrast to the U.S. and the UK, there are no intermediate colleges between high school and university. The UK and U.S. schooling systems are more similar: children start school at the age of 5, and complete compulsory schooling after 11 years; in both cases there are intermediate educational levels between high school and college.

In order to keep the analysis as consistent as possible, the classification criterion applied is the highest educational qualification which is common to all countries and whose information is available in all datasets.

Therefore the three educational groups are defined as follows:

1) Lower;

2) Intermediate;

3) High.

The lower education group includes workers who have completed compulsory qualifications i.e. less than a lower secondary education; this corresponds to 8 years of completed education in Italy (which also includes individuals with no education at all); for the U.S. the same group corresponds to 0 to 11 years of schooling and for the UK this aggregates individuals from 0 to11 years of schooling. The intermediate category gathers workers with qualifications that are between high school dropout and the degree (both excluded). In Italy, this corresponds to workers with a high school "diploma" (13 years of schooling); in the UK and the U.S. this corresponds to any qualification with years of schooling equal or greater than 12 and less or equal to 15 years of schooling. The educational group "high", in all cases, refers either to graduate or postgraduate earned qualifications.

[19] The actual university system is more similar to the English and American ones; university is composed of two steps: three year courses plus 2 year courses which are not compulsory. Data contained in the Bank of Italy all refers to the pre-reform system i.e. 4 year courses.

The variable years of potential labour market experience are conventionally derived as Age – Years of completed Education- the age at which children start school in the three countries. Workers are aggregated into four-year experience intervals (1 –10; 11-20; 21-30, 31+). Based on the three education categories (lower, intermediate and high) and the four experience categories, workers can be classified into one of 12 skill groups[20].

In doing so, this study uses a lower number of dummies for years of completed education than Lemiuex does, but this has the advantage of being consistent through data sets[21]. Lemieux uses CPS earnings weights together with the number of hours worked per week as weight in the overall analysis. However since weighting makes very little difference to the US data, all results are based on unweighted data. The data appendix provides additional details of three datasets used and in particular discusses the issues regarding education variables.

The residual distribution measures (captures) the distribution and the effects of unobservable skill components that are related to the specific educational group to which workers have been assigned to.

Considering a simple wage equation such as (1), the wage equation residual ε_{it} can be conceptualizes as having two components: an individual's percentile in the wage distribution η_{it} and the distribution function of the residuals F_t ( ). By the definition of the cumulative distribution function, we can write the residual as:

(1a) $\varepsilon_{it} = F_t^{-1} (\eta_{it} | x_{it})$,

Where the $F_t^{-1} (. | x_{it})$ is the inverse cumulative residual distribution for workers with characteristics x_{it} (education and experience) in year *t*. The equation (1a) allows to captures the changes in the distribution of the residuals, that are linked to the different educational attainments of workers, to the total changes in inequality.

The residual distribution depends on the (x_{it}) characteristics of the workers, e.g. depends on the educational attainments of workers, therefore the residual inequality analysis might be sensitive to the educational groups to which workers have been assigned to as well as to the criterion adopted to define the groups.

[20] In the original data, Lemieux defines 20 education and experience groups based on 5 education groups and 4 experience groups. The education groups are high school dropouts; college; some college; post graduate and high school college. These 5 groups have been clustered into 3 to generate 3 educational groups consistent through the three countries.

[21] The categories used for years of schooling completed are: 8, 13, 17 for Italy; 11, 13, 15 and 17 both for the U.S. and the UK.

Because the empirical analysis of this chapter uses data from countries characterized by different educational systems, in order to make the analysis comparable the educational criterion adopted is the highest level of education. However, this criterion might of course involve measurement errors in the estimates depending on how workers have been classified into the three levels of education in the three different countries. For example the "lower" level of education classifies workers with 11 years of schooling in the UK and the U.S. while this corresponds to 8 years of schooling for Italy, meaning the lower level of education criterion compares workers that have different years of schooling and this would possibly reflect on the observed trends and levels of residual inequality.

For example one of the common educational criterion used for the UK literature is age left full time education; the lower level of education would correspond to the group of workers who have left school at 16 and under, that will include those with GCSEs as well as those who had 2 years of post compulsory education for the older cohorts. As a consequence the educational group might not fully reflect changes in the dispersion of the residual, and also for those workers with more education in the same group, cohort and age effects are, if anything, mitigating the increase in residual wage dispersion.

As pointed out by Gosling et al. (2000) there are few sources of increasing heterogeneity within each education group; one is the set of qualifications received within each of the groups based on the years of education measure; another is the possible changes in the distribution of other skills across cohorts and the last is changes in the composition of ability in each group solely driven by the fact that the process determining educational choices may have changed. They provide evidence that the within group wage dispersion grows more or less the same rate across education groups when these are defined by years of schooling; the overall increase in within group inequality is smaller when defined by educational qualification. More interestingly they show that using years of education rather than qualifications obscures the fact that there has been no increase in within group inequality amongst the group of workers with a degree. Similarly the lower educational group displays a higher increase in within group inequality when using the years of education measure.

This has to be borne in mind when interpreting the results and comparing the trends cross countries.

2.6. Results

2.6.1 Descriptive Results

Figures 2.1a to 2.2.b present the trends in log hourly wage inequality from 1987 to 2004 using the CPS MAY/ORG, GHS and SHIW samples. Following the main approaches of literature (JMP 1993; Katz and Autor 1999, Lemieux 2006), the summary measures of wage dispersion used are the standard deviation of and the 90-10, 90-50 and 50-10 percentile differences in log hourly wages.

Figures 2.1a and 2.1b plot the standard deviations for men and women in Italy, the UK and the U.S. The figures display some remarkable differences among the three countries both in levels and occasionally in the direction of the trends. The level of wage inequality in the U.S. remains the highest of the three countries considered for the whole period. CPS data shows that overall inequality for male wages, on this measure, in the U.S. remained relatively flat through the sample years compared to the other countries, but rose among women, with the main increase occurring at the end of the 1990s. Although the level of wage dispersion is pronounced for both men and women, wage dispersion for male workers is slightly higher than for women. In 2003 the standard deviation of hourly wage was 0.582 for men and 0.541 for women. More variation characterises the trend over time for women compared to men, with the bulk of the increase occurring during the 1990s. Female workers also exhibit a higher increase in wage dispersion: the change of standard deviation from 1987 to 2003 was about 0.012 for males and 0.032 for females; this change appears to be smaller if compared to the variation which occurred during the 1980s; from 1980 to 1990 the change in standard deviation was 0.0724 for males and 0.102 for female workers.

Some interesting features characterize the trends in hourly wage inequality in the UK. First of all, the level of inequality is quite high and, for men but not women, closer to that of the U.S. than to Italy. Similarly to the U.S., the UK gender gap in inequality appears to be an important element of the inequality story since the standard deviation of log wages is higher for men than for women in the US and the UK but not in Italy. During the 1990s, there is no obvious trend in inequality for men; wage inequality increases both for men and women until 1991; and then slows down, particularly for women, after 1991; rises in the

middle of the 1990s and then falls after 2000. The decrease is more evident for women.

The main message of Figures 2.1a and 2.1b is that the evolution of wage inequality in the UK differs from the other two countries, with the UK being the only country to experience an overall decrease in wage inequality over the sample period, especially for women. In fact between 1987 and 2003, the standard deviation of log hourly wage decreases by 0.011 for male workers and 0.055 for female workers.

The fall in wage inequality in the UK at the end of the 1990s may be explained, in part, by the introduction of the minimum wage. Empirical evidence provided, amongst others, by Gosling and Lemieux (2004) documents that the decline in wage dispersion during the same period is due to changes in labour market in particular to the decline of unionization and to the introduction of the minimum wage[22] in the UK at the end of the 1990s.

Moving to Italy, the figure displays two clear features: wage inequality affects women more than men and the large rise in wage inequality occurring in the early '90s is not seen elsewhere in any other period or any other country. From 1989 to 1993 wage dispersion increased by 7 percentage points for men, and 12 percentage points for women. Italian women also experienced the highest change over the whole sample period, with the standard deviation increasing by 0.041. Comparing male workers in the U.S. and in Italy, more similarity in the trend arises from 1993 to 2001 when, in both countries, wage dispersion remains relatively flat, though levels remain different.

Figures 2.2a and 2.2b contrast the evolution in log hourly wage percentiles both for men and women from 1987 to 2004. The graphs plot the trends of the 90-10 percentile gap as well as the trends in the upper and lower tails of the wage distribution. The results broadly confirm the well documented story about changes in wage structure between workers located in different parts of the wage distribution (Katz and Autor, 1999; Autor, Katz and Kearney, 2008; JMP 1993; Lemieux 2008).

[22] Gosling and Lemieux analyse the relationship between the minimum wage and wage inequality; and show that the fall of the real value of the minimum wage contributed to the rapid increase of wage inequality in the U.S. during the 1980s.

Workers in the upper percentile (90) of the wage distribution gain consistently on workers in the lower percentile (10), with resulting increasing inequality; this pattern is particularly evident for both men and women in the U.S. while the gain is much less pronounced for both Italian men and women workers. More fluctuation characterizes the 90-10 gap trend for the UK, especially for male workers.

The 50-10 wage gap stopped rising after the mid-1990s in Italy and the UK and this phenomenon was greater for men. Between 1997 and 2003 in the UK the 50-10 gap for men and women decreased respectively by 0.089 and 0.031 log points. For female workers in the U.S., the lower tail gap gradually fell by 0.036 log points with the decrease being larger for men (0.171). For Italy, there is more compression in the lower tail inequality, moreover between 1987 and 2004 there is almost 0 change for women and an increase of 0.033 log points for men.

By contrast, upper tail (90-50 gap) wage inequality increased steadily in both the US and the UK over the sample period and was more pronounced after 1995; whereas in Italy the upper-tail wage inequality decreased, particularly after 1995.

These log hourly wage differentials plot to some extent similar patterns of inequality for men and women in all countries. The divergence in the upper tail inequality is remarkable in Italy and even greater for female workers in Italy and the U.S. Again, fluctuations in trends characterise inequality in wages for English male workers.

The diverging path between the 90-10 log wage differential and the 50-10 and 90-50 ratios for both men and women is most marked for the U.S., while for Italy the divergence between the earnings of more skilled and less skilled workers is less evident. The rise in the upper tail earning inequality accounts for most of the inequality in all countries, both for male and female.

One of the aims of this analysis is to explain how changing age and educational composition of the labour force affects wage inequality. Table 2.1 provides comparisons of educational level and experience of workers. The trends in the share of the educational level of the work force are relevant in this analysis to compare later with the composition effects. In particular, if the composition hypothesis stated by Lemieux holds, we should expect composition effects to be

more marked in the period in which the educational achievement of the work force changes.

The first row of each panel reports the share of workers, separated by gender, with respect to earned educational qualifications. The striking aspect is the difference in the shares of lower educated workers between the U.S. and the other two countries. In all cases, the share of less educated workers decreased through the years. The percentage of workers with low education in the U.S. is small compared to the UK and Italy; in 1987 only 18 % of the male workers in the CPS data had a low education (14 % for women); the highest percentages of lower educated workers were in the UK GHS sample (63 % of male workers and 69 % female in 1987). The Italian experience of the lower educated workers is similar to the UK. At the beginning of the sample period, 58 % of male workers (41% for women) in the SHIW data only had a low level of education.

The educational performance of Italian workers is between that of the U.S and the UK. The full time employment share of workers with a degree is around 30% in 2003 in the CPS. Almost no change occurs in the degree level for Italian male workers, while the high level of education for Italian female workers in the SHIW is nearly double that of men in all years. In the UK, the variation in the share of higher educated workers is characterized by the fact that in the last 16 years, female educational attainment overtakes that of men. The share of higher educated workers increases over time; in 2003 the share of male high educated workers is 22% compared to the 13% of the 1987; for the female workers the increase is even more striking: from 1987 to 2003 the share of women in the labour force with a graduate or postgraduate degree, increases by 17 percentage points. These differences in the educational level of the workers in the samples could therefore influence residual inequality.

Differences also arise with respect to the level of experience. The U.S. had the highest share of low experience workers in 1987 and the lowest share of workers (16%) with more than 30 years of experience. Italy had the lowest share of younger cohorts of workers in 1987 and the highest share of more experienced (older) workers, about 30% for men. Differences arise when comparing male and female workers, particularly in the UK and Italy. In 1987, 47% of English female workers belonged to the new entry cohorts. In Italy this was 25%, 10 percent more

than the equivalent estimate for men, suggesting an increase in labour force participation of women in both countries.

The analysis of trends and levels in wage inequality for the UK, U.S. and Italy cannot exclude an analysis of levels and trends in employment rates. These are plotted in Figure 2.3a for men and Figure 2.3b for women. The existence of the dichotomy between the U.S. (and the UK) experiencing robust employment growth and widening inequality, and Europe (Italy) where the index of wage inequality showed little or no change and employment was stagnant, is well known in the economic literature and discussed by Krugman (1994).

Both figures are based on the Organisation for Economic Co-operation and Development (OECD) data from 1987 to 2004 and document the existence of an employment gap between the UK, the U.S and Italy. In both the UK and the U.S., for male workers, the employment rates during the sample period analysed, subject to cyclical variation, were around 80% and were very similar between the two countries, with the U.S. displaying a higher employment rate between 1991 and 2001. Italian male workers experienced a lower employment rate ranging from 67% to 70% during the same time period analysed.

The employment gap between the U.S, the UK and Italy is striking for women, in fact while the employment rate for women ranges from 58% to 67% for the UK and 62% to 68% for the U.S., Italian women faced a much lower employment rate (35%) at the beginning of the time period analysed. Although this had increased by the end of the 1990s, reaching 45% in 2004, and converged toward that of the US and the UK, Italy remains one of the EU countries with the lowest female employment rate.

2.6.2 Residual Wage Inequality: Convergence or Divergence Patterns?

The next results provide evidence of trends in residual wage inequality for the U.S., the UK and Italy, starting with a description and comparison of trends and levels of between and within variance.

Figures 2.4a and 2.4b plot the trends for the (actual) residual (within-group) variance and the between group variance respectively for men and women from 1987 to 2003 for the U.S. and the UK and from 1987 to 2004 for Italy, based on a decomposition of equation 6.

The figures help investigate to what extent log hourly wage components (within-group and between) affected the overall increase in wage dispersion in the U.S., Italy and the UK during the last 16 years. Comparing the patterns for the three countries, consistent with the descriptive analysis of part one, the U.S. has the highest level of within-group (residual) variance for both men and women. Most of the increase occurs during the 1990s, decreases slightly at the end of the 1990s and rises again between 1999 and 2003. The change in prices of observables, represented in the graph by the between variance, remains relatively flat during the period analysed. More variation occurs for observable skill prices for both Italy and the UK which do not correspond to a similar pattern for the unobservable skill variance. For Italy, from 1993 the between variance decreased while the within increased, this appears to be true for both men and women. The between group variance also decreases for both men and women in the UK.

These findings validate the results initially suggested by Lemieux for the US: residual inequality appears to have an increasing role in explaining the overall increase in wage inequality, and this holds for all cases. Some similarity can be noticed with respect to the evolution of the between variance: the role of the observable components, in this context education and age, appears to lose power in explaining wage inequality particularly after the middle of the 1990s mainly for Italy and UK.

Figure 2.4b shows that for women in the UK between 1987 and 2003 the between-group variance decreases (-0.051) more than the within-group (-0.001).

For Italian male workers, the between group and actual residual variances move simultaneously during the initial period; but after 1993 the variance of unobserved skill increases steadily and decreases just after 2002, while the between goes in the opposite direction.

Different patterns and timing in the changes of the unobserved and observed skill variances for the U.S., the UK and Italy suggest that the returns to observed and unobserved skills may differ. When the residual variance grows more than the between group variance this suggests that unobserved skills play a more important role than observed skills in shaping the overall wage dispersion. In other words, similar to the U.S. within-group is majority of total variance, unlike the U.S., in both the UK and Italy the between group level (and share) is falling.

To complement this result, tables 2.2a and 2.2b report the R^2 of the yearly OLS regression, which represents the contribution of the between components in explaining the changes in log hourly wages based on the above model (3). The values range on average between 0.32 and 0.39; differences arise for the U.S. between men and women, in fact the values of the R^2 of the OLS regressions for women are on average 0.05 lower than that for men. For the UK, there is not much difference between men and women and similarly the values of the R^2 decrease over the sample period. For Italy, the R^2 explains a little bit more for women (0.39) than for men (0.36) and similarly to the UK the between component looses part of the explanatory power over the sample period.

Tables 2.3a, 2.3b and 2.3c summarize the growth in the residual variance relative to the growth in the total variance for the U.S., the UK and Italy, separately for men and women. Table 2.3a shows that for the U.S. from 1987 to 2003, changes in the residual for men accounted for 86 percent of the overall change in the total variance (53 percent for women), suggesting that more than half of the growth in wage variance is due to the residual component. This table also documents that holding the distribution of skills at its 1987 level, the remaining growth in the residual variance accounts for a small negative percentage (-0.01 for men and -0.16 for women). The results do not qualitatively change when holding characteristics of workers at the 2003 level; again, only 18 and 13 percent respectively for men and women, is explained by residual variance; this suggests that had the level of education remained constant at its 1987 level, the residual variance would account for only a small fraction of the overall rise in wage inequality i.e. the great part of the variation in the actual variance is due to the fact that the workforce is more educated and more experienced than it was in 1987.

Table 2.3b reports the changes in residual variance with respect to changes to overall wage variance for the UK from 1987 to 2003, separately for men and women. The main feature for the UK is that, in contrast to both the US and Italy, the total variance in the period analysed decreases both for men and particularly for women (0.011 for men and 0.053 for women); the decrease starts at the beginning of the 1990s. The main divergent pattern between men and women is that despite the fact that women experience a decrease in both total variance and residual wage, although with different magnitude, for the English male the

decrease of total wage variance contrasts with a higher increases of the residual variance. Holding the skills distribution at its 1987 level, only 17 % of the total wage variance is explained by the residual variance for men and 12 % contribute to increase the variance of wage. The larger decrease in wage variance occurs during the 1990s.

Table 2.3c summarises the same results for Italy from 1987 to 2004; in the overall period the variance of wage increases both for men and women, with a more pronounced increase for men; in this case the within-group variance accounts for almost all the variation, and even when controlling for distribution of skills, both in 1987 and in 2004, the residual variance still explains the bulk of the increase suggesting a very marginal role for the composition effects. For the female Italian workers, the increase of the residual variance is even larger than the increase of the overall wage variance itself.

To better understand the link between changes in shares of unobservable skills and changes in the education and experience characteristics of the labour force and therefore to separate the effect of skill prices from that of composition, tables 2.4a-2.4f explore the role of the work force composition. The tables report the value of the within variance for 12 education and experience groups[23]. To obtain larger cells and therefore to improve the precision of the estimates, two years at the beginning of the period (1987 and 1989)[24] and two years of the end of the period (respectively 2002 and 2003 for the U.S. and the UK; 2002 and 2004 for Italy) have been merged together.

Tables 2.4a and 2.4b illustrate the evolution of the residual variance for the U.S. The results in column 3 show that the change in the residual variance is always negative for all lower educated groups, for both men and women. For men, the variance declines over the two periods considered for the younger cohort of the intermediate group, while remains relatively unchanged for more experienced workers of the same educational group. Similarly, in the same group for women, within variance decreases for the two younger groups and increases for the older intermediate groups. In all cases considered, the decrease of the variance corresponds to a decrease in the corresponding work force share, as shown in

[23] Lemieux's original work derives 20 groups; the data appendix motivates in detail this decision.
[24] The choice of these years is based on the availability of the point in time years of Italy, as better explained in the data appendix.

columns 3 and 6. This finding is consistent in part with one of the main findings of Lemieux for the U.S.

High educated workers instead experience a positive variance change, which is confirmed for men and women, though the level of the residual variance for men is larger than the residual variance for women.

Tables 2.4c and 2.4d report trends in residual inequality by education and experience group for the UK for men and women respectively. No systematic or uniform pattern emerges from the GHS dataset. The increase (decrease) in residual variance is not always associated with an increase (decrease) in the corresponding share of group workers. Only for higher educated male workers there is a clearer pattern: the positive change in the share of more educated workers corresponds to an increase in the within variance. More uniform trends appear for the shares of education and experience group in the workforce presented in column 6. Trends are similar across groups and between men and women; the share of the intermediate group always increases except for the younger group; at the same time the share of higher educated workers increases. The changes in the trends are quite similar for female workers, although the increase in intermediate and higher educated group share is more rapid.

Tables 2.4e and 2.4f give the results for Italy. Column 3 of table 2.4e displays changes in the within-group variance for male workers. The clear pattern emerging from the analysis of male Italian workers is that, unlike the U.S., the residual variance increases for all education and experience groups. The highest value occurs for the youngest, more educated workers in the sample; changes in the workforce share are decreasing for the lower educated, and increasing for each and every experience group for the intermediate. Changes for high educated are diverging, in fact there is a decline in the share of younger higher educated workers and a slight increase for the two older groups.

The results for women are qualitatively similar, with the exception of one old intermediate group experiencing a small decrease in within variance. The trend for the change in share of work force is the same for lower educated female groups, similar for the intermediate, with the exception of the younger group that declines; and the share of higher educated women increases more than that of men.

To summarise, Italy differs from the U.S.; despite similarity in the decrease in low educated share, residual variance has increased, while the UK lies in between. There are two main reasons which may explain the differences in the role and trend of residual with respect to the overall wage inequality in the three countries. One reason that may explain cross-country differences in the residual variance trends can be linked to the differences in the institutional labour market structures. Autor and Katz (1999) explain that the same underlying demand and supply stocks may have differential effects on wages and related components, depending on differences in wage-setting and other labour market institutions. The stronger the role of labour market institutions and the less responsive the institutions are to changes in market forces, the greater the impact is likely to be on employment rather than on wages. This may be the case with Italy, where the existence of the collective bargaining agreement, as well as the agreement negotiated at national level between unions and employer associations, makes the labour market less responsive to changes in composition of the labour force and therefore prevents the reward of both observable and unobservable skills. In fact, as shown above in the paper, the return to observable (education) in Italy is lower than the return to education in the U.S. and the UK.

The other reason is that changes in supply characteristics are less dramatic in Italy compared to the U.S. and the UK; for example in Italy, the share of higher educated workers is lower compared to both the U.S. and the UK. As pointed out by Natticchioni, Ricci and Rustichelli (2008), in Italy the share of individuals who had achieved a degree in 2002 was 10%, highlighting a lower educational attainment than other European countries. The striking result for Italy, therefore, is that even though the share of the lower educated work force decreases over the years, as in the U.S., residual inequality of less educated workers increases rather than decreases.

This suggests that in the Italian context, Lemieux's composition effects do not exert the same mechanical force with respect to changes of the residual: the increase of residual variance does not appear to be driven by changes in the composition of the work force. This result, confirmed by the findings displayed in table 2.1c, may also depend on the fact that changes in educational level of the labour force have been more modest compared to the other two countries, as

shown in table 2.1, particularly when considering the share of workers with a high educational level.

A clearer picture of the evolution of the residual variance by educational composition of the work force is provided in figures 2.5a and 2.5b.

The figures plot the within variance year-by-year, calculated as an average across experience groups for each educational group, separately for men and women. The trends are different for each country and across different decades.

For the U.S., the trends in within variance for the three different educational groups are divergent for both men and particularly for women. In both cases, the level of the variance for more educated workers is the highest and is steadily increasing over time. Less change occurs for the intermediate group, while for the less educated, the residual variance declines over time.

Although the trends in variance by education groups in Italy are lower than in the UK and the U.S., some similarities with U.S. wage inequality timing can be observed. The SHIW data displays the higher and increasing level for the residual of the higher educated group; less variation appears for the intermediate group but, differently from the US, the variance for the lower educated group grows in tandem with that of the higher educated group, both for males and females, particularly after 1990.

For women, the trend for more educated follows that of men. However, what is particularly interesting for the Italian women is that in contrast to men, starting from 1997 the lower educated group variance fell.

The results for the UK are difficult to interpret because there is no uniform pattern dominating the whole trend for either male or female workers. The clearer pattern is displayed for the within variance of less educated workers; the level is always lower than the other two educational groups. The UK trends go somewhat against the SBTC hypothesis since the SBTC would predict a continuing and steep growth in wage inequality throughout the 1990s.

Figures 2.6a and 2.6b present the results of the counterfactual reweighting approach proposed by Lemieux that helps to account for the role of composition effects on changes in residual wage inequality. The figures compare for each country, separately for men and women, the actual residual variance from 1987 to 2004, to the counterfactual variance that would have existed if the distribution of

skills had remained at the same level of the initial or the final year. The composition effect is represented in the figures by the distance between the actual and the counterfactual variance.

Comparing the role of changes in the characteristics of the labour force in the three countries, the composition effects are different across the three countries, suggesting that changes in the shares and levels of education and experience of the work force had a different impact on the evolution of wage inequality and particularly on the evolution (rise) of residual wage inequality.

When the distribution of skills is held at its 1987 level, the composition effect appears to play a central role for the U.S. for which the great part of the growth in residual can be explained by the changes in education and experience composition. For Italy, changes in the composition of the labour force do not seem to be important to the increase in residual inequality, having a small, negligible role since 1995. The results are similar for women, except that the composition effects play a small role during the last years of the samples both for Italy and the UK. The effects for US women are more alike in terms of timing to that of men. The results are consistent with Lemieux, confirming that as the US labour force has become more educated and older, this also increases the role of unobservable skills of workers. If the workforce had 1987 characteristics, the level of the residual would have been lower, suggesting that composition effects exerted an upward force on residual inequality. This is less evident for the UK, except when considering the upper and lower residual gap trends.

2.6.3 Residual Inequality: Alternative Measures

The remaining figures (2.7a-2.9b) look at alternative measures of wage inequality, specifically the 90-10, 90-50 and 50-10 residual gaps, comparing the actual residual and the residual computed holding distribution of skills at their 1987 characteristics, to understand how much change in residual, as well as composition, affects workers at the different tails of the distribution.

The 90-10 gap, plotted in figures 2.7a and 2.7b for men and women respectively, has been increasing steadily through time for men in the U.S. For women, the 90-10 gap increases until 1995, decreases during the latter part of the

1990s and increases faster thereafter. Holding the characteristics of workers at their 1987 level, a clear effect of composition changes is displayed for the U.S. Minor composition effects can be noted for Italy, both for men and women; and for the UK, the trend is similar to that for men, the composition effect starts to become evident to some extent after 1995; the pattern reverses and becomes unclear for English women.

One interesting result arises in the graphs of the evolution of the 50-10 gap (2.8a and 2.8b); the patterns for the three countries are quite different. For the U.S., when the characteristics of education and experience of the labour force are held constant, the effect of changes in labour force is clear and more striking for women. For Italy, contrary to the results for AKK (2005) in the U.S., changes in education and experience composition have no effect for workers located in the bottom half of the distribution, for which changes in wage inequality appear to be due to reasons other than the changes in work force characteristics. Again, no clear or uniform pattern can be observed over time for the less skilled UK workers.

For the U.S. the trend is qualitatively similar between men and women, in fact the 50-10 gap declines more consistently through time, and slightly increases after 2001. The decrease during the 1990s is more pronounced. In both cases, when holding distribution of skills of workers at their 1987 level, the level of the 50-10 residual variance gap is lower, suggesting the existence of composition effect on workers located at the lower part of the wage distribution.

The 50-10 residual gap for the UK displays more variation both for men and women, with a level of residual gap lower, starting from 1997. Composition effects appear to keep the level of residual lower between 1989 and 1998 for women, with no remarkable effect during the last years. By contrast, no effects can be noted for men, except a small effect since the middle of the 1990s.

More similarity can be noticed when analysing the trend and timing of the residual upper tail (90-50) distribution plotted in figures 2.9a and 2.9b. In the United States, the residual inequality of the 90-50 gap increases over time and is affected by changes in the work force, both for men and women, even though a stronger effect is displayed for women. The Italian trend of the 90-50 residual gap is more similar to the U.S. than it is to the UK. For Italy, composition effect can indeed explain part of the increase in the 90-50 residual gap, both for men and

women. The fact that composition effects affect differently the residual of the lower and upper tail of the wage distribution, recalls the importance of how the rising skill prices reflect differently on high and less skilled workers.

These results confirm one of the main stylised facts about wage dispersion: the increase in the 90-50 gap; and this holds for all countries.

More variation occurs for the UK, and for all countries holding distribution of skills at their 1987 level can help to explain the increase in the upper tail of the residual.

2.7 Conclusion

This chapter presents a comparative analysis of the evolution in residual hourly wage inequality for men and women in the United States, the United Kingdom and Italy between 1987 and 2004, using microdata from the Current Population Survey for the U.S., the General Household Survey for the UK and the Survey of Household and Income Wealth for Italy.

The analysis shows some remarkable differences among the three countries both in the level and sometimes in the direction of the trends of wage inequality over the sample period, as measured by the standard deviation of log hourly wage. The level of wage inequality in the U.S. is the highest amongst all the countries and throughout the years, even though it remains relatively flat through the period compared to the other countries. The level of inequality in the UK is quite high and also higher than that of both the UK and Italy. The UK experience, in terms of wage inequality, differs in this context from Italy and the U.S. since it has been the only one to face an overall decrease in wage inequality; from 1987 to 2003 standard deviation of log hourly wage decreases by 0.001 for male workers and 0.053 for female workers. Two features characterize wage dispersion in Italy: the highest change of wage inequality that occurs in the early ’90s is not seen elsewhere in any other period; Italian female workers are more affected than males, from 1989 to 1993 wage dispersion increases by 7 percentage points for men, and by 12 percentage points for women.

The results provide some support for conclusions concerning residual inequality trends proposed by the main empirical literature (Juhn, Murphy, Pierce 1993, Autor, Katz and Kearney 2005, 2007; Acemoglu 2002, Lemieux 2006): the

residual accounts for most of the wage inequality variation in all three countries (though of course this in part reflects the amount of disaggregation in the observed model covariates). In the U.S. from 1987 to 2003, changes in the residual for men account for 86% of the overall change in the total variance, and 53% for woman. In the UK, despite the decrease in overall variance of wage both for men (0.033) and women (0.034), the residual variance increased by 0.014 log points. More strikingly, the within-group variance in Italy corresponds to the wage variation, explaining almost all the change in wage inequality, with the bulk of the increase occurring during the 1990s, confirming the main results of existing literature in Italy.

Comparing the role of changes in characteristics of the workforce in the three countries, composition effects are different among the three countries, suggesting that changes in the share and level of education and experience of the workforce had a different impact on the evolution of wage inequality and particularly on the evolution (rise) of residual wage inequality reflecting the fact that the education/experience shares grew differently across countries, therefore shaping differently demand and supply of observable and unobservable skills. The composition effects hypothesis appears to hold only for the U.S.; in the UK and Italy, when controlling for distribution of skills, both at 1987 and at 2004, the residual variance still explains the bulk of the increase, suggesting a very marginal role for the composition effects.

In partial contrast with AAK (2005) who show that composition effect was concentrated in the lower tail of the earnings distribution, results for Italy show that compositional changes only affect the upper tail (90-50 gap) earnings distribution.

Appendix A2.1: Data

United States:

Since Lemieux provided data for MAY/ORG CPS, the log of hourly wages has already been defined. As Lemieux explains, for 60 percent of the work force in the sample, information about the hourly wage has been directly collected from workers; for the rest of the U.S. sample, log wage has been computed[25].

In the whole analysis, Lemieux uses as weight hours worked per week, together with CPS's weight; after several iterations consisting in weighting and unweighting the analysis, using, for example, the weight defined as: "hours worked per week/10". Because of negligible differences in the estimates, weights have not been used in any data country. This decision has also been supported by the fact that the GHS and the SHIW do not contain any earning weights which may be compared to those available in the CPS.

The only weight used is the counterfactual one defined as follows:

$$\text{Weight} = (1-\omega)/\omega$$

where ω is computed as in Lemieux's by using the logistic regression.

The original work defines twenty education and experience groups, which are based on four experience categories (0-10, 11-20, 21-30 and over 31 years of experience) and on five education categories (dropout, high school graduate, some college, college graduates, college post graduates). The regressions use 9 education categories (0-4, 5-8, 10, 11, 12, 13-15, 16 and 17+).

United Kingdom:

Part of the empirical work on wage inequality in the UK uses an alternative data set: the Labour Force Survey (LFS). The LFS, like the GHS, contains information on individual wages and employment status, and is a larger sample, but data on wages are only available from 1993.

The main advantage in using the GHS is that it contains information in all periods considered. The main disadvantages lie in the sample size, in particular,

[25] For details see Appendix A in Lemieux 2006.

because the educational level is not recorded consistently through time. The GHS contains information on the highest level of qualifications as well as on the age left school. The main limitation, which is reflected on the other datasets, is that from 1975 to 1996 there is no separate information between post-graduate and degree; in the last years of the sample, education categories are nearly 40. Real wages are obtained by deflating nominal wages with the Retail Price Index (RPI) based at the 2003 level.

Italy:

The SHIW covers the years from 1977 to 2004. Since information on the hours worked per week is only recorded starting from 1987, it has been necessary to start all periods from 1987. The survey is run every two years, therefore for Italy the mid-points of the time intervals available from the Bank of Italy and used in this research are: 1987, 1989, 1991, 1993, 1995, 1998, 2000, 2002 and 2004. Moreover, for the years before 1984, information on age is only available in class and is not continuous.

Information about hours of work also includes the overtime hours (paid). In the SHIW there is separate information about hours worked per week and overtime with corresponding wages. Because all data sets include overtime in hours worked, observations with overtime worked have been used in the SHIW data.

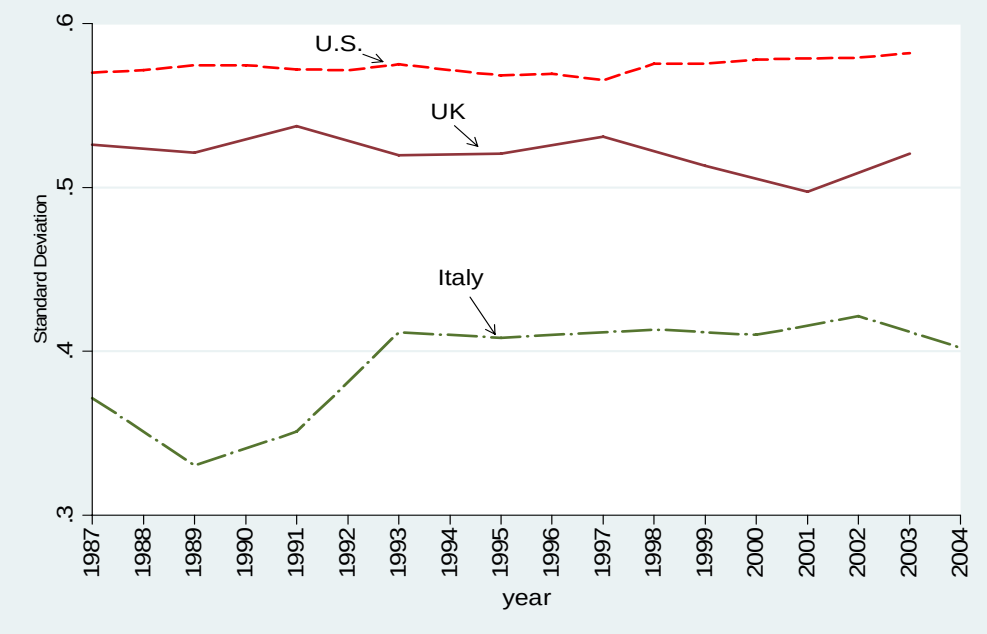
Part of the existing literature on earnings inequality in Italy uses an alternative data set available, the Italian Institute for National Social Security (INPS). The INPS is an administrative data set that, based on firms' declarations, collects data on all Italian workers employed in the private sector with respect to the period from 1985 to 1999. For each calendar year, the Social Security forms of employees born on the 10^{th} of March, June, September and December were selected to form random samples of population of employees.

The main advantages of using INPS data are that it contains annual gross wages and that the sample size is larger than the SHIW one. In fact, each yearly sample includes approximately 100,000 workers of Italian private firms, excluding agriculture and central state administration.

Unfortunately, the INPS does not contain any information on hours of work per week, on workers in the public sector or on the educational qualifications of the workers; information which is indeed crucial in this context.

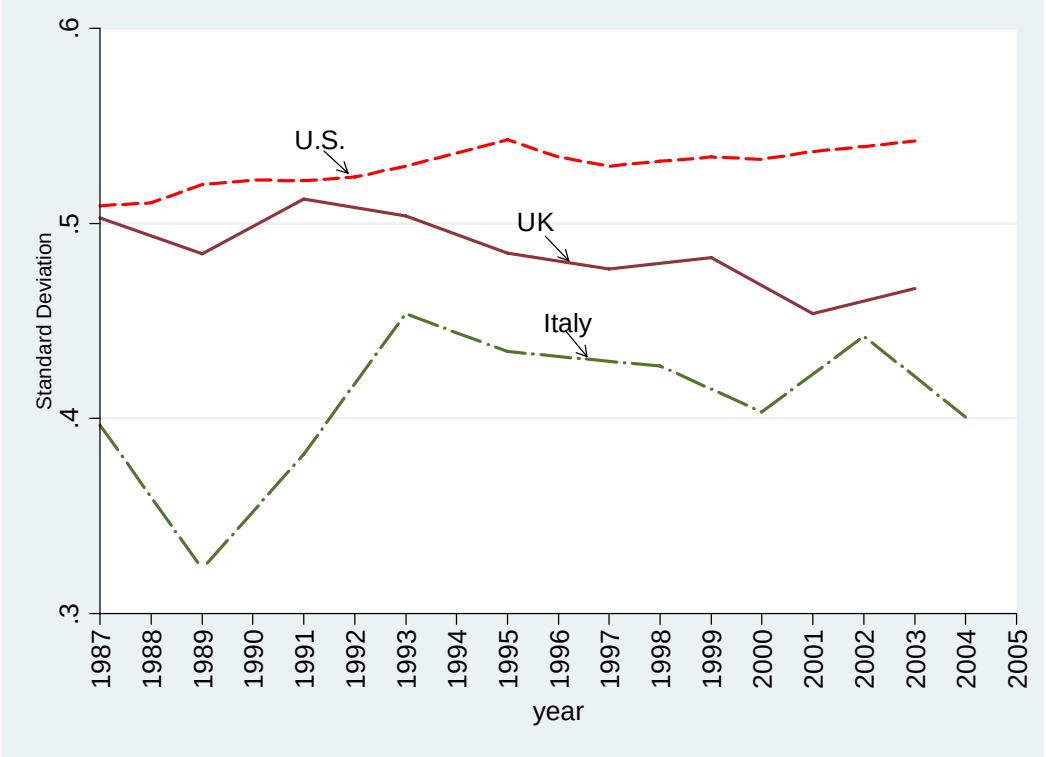
For the SHIW dataset, wage values that are between 1 and 50 Euros (at 2004 Euro value) are used, observations deleted in this way are very few (9 for hourly wages greater than €50, and 78 for hourly wages less than €1).

Figure 2.1a: Standard Deviation of Log Hourly Wage, Men



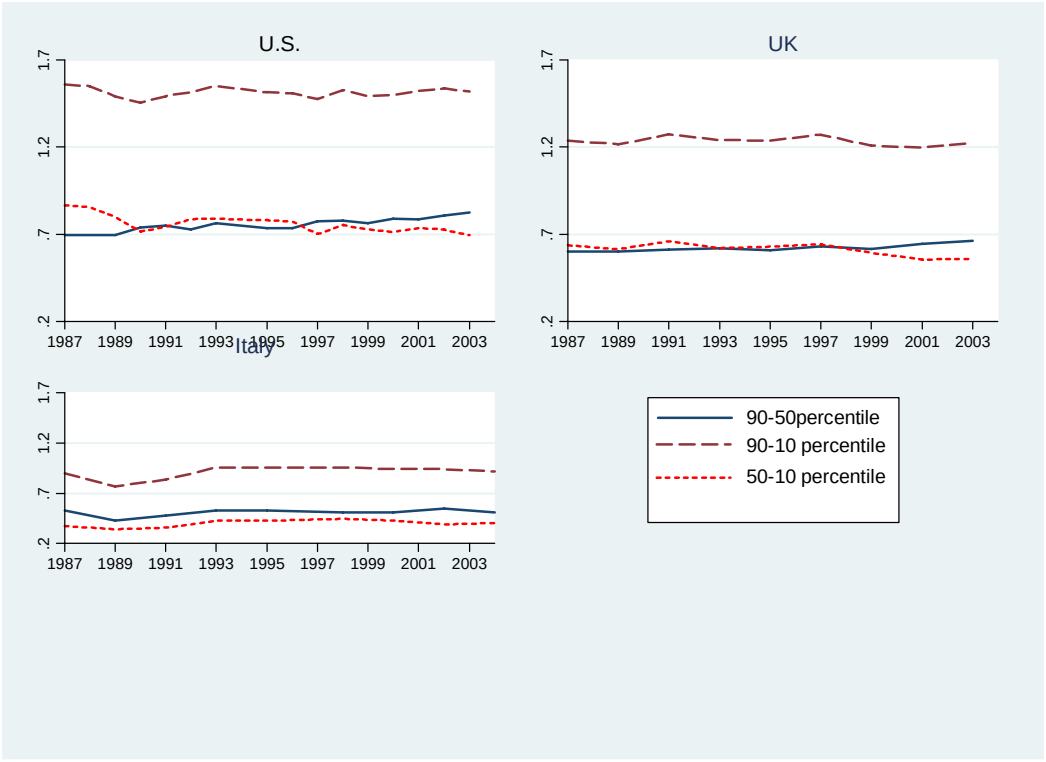
Notes: Based on CPS, GHS and SHIW. Samples include male workers in labour force, employed only, working full time, considering main job with positive potential experience.

Figure 2.1b: Standard Deviation of Log Hourly Wage, Women



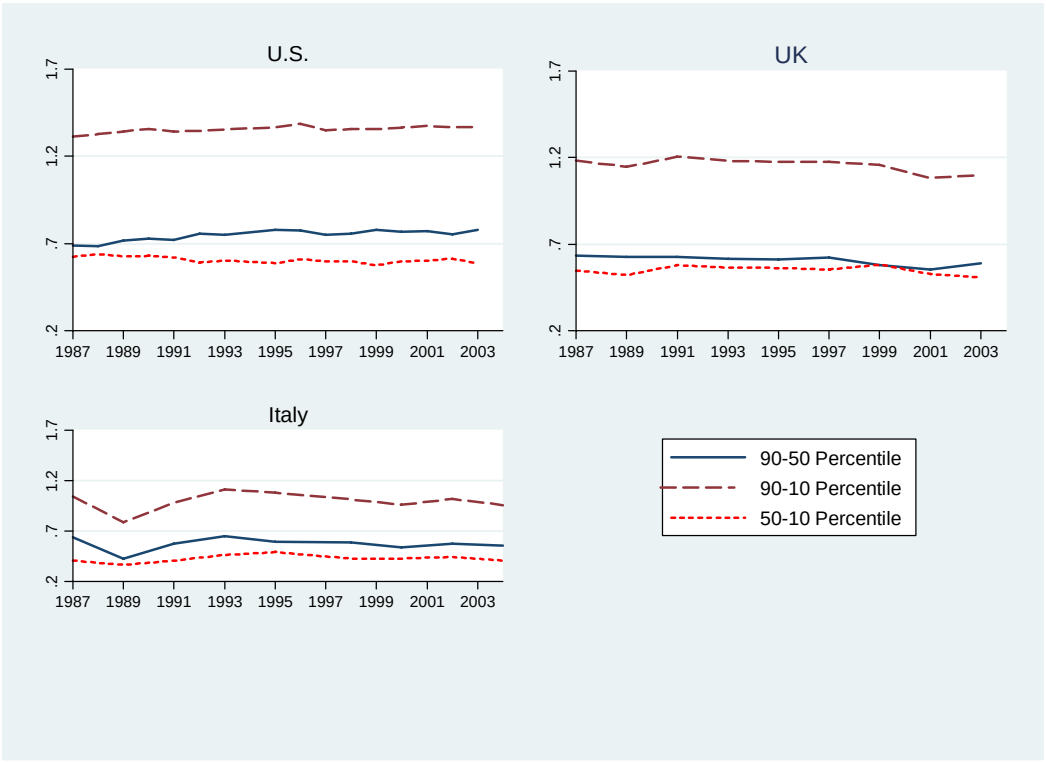
Notes: Based on CPS, GHS and SHIW. Samples include female workers in labour force, employed only, working full time, considering main job with positive potential experience

Figure 2.2a: Percentile of Log Hourly Wage Distribution, Men



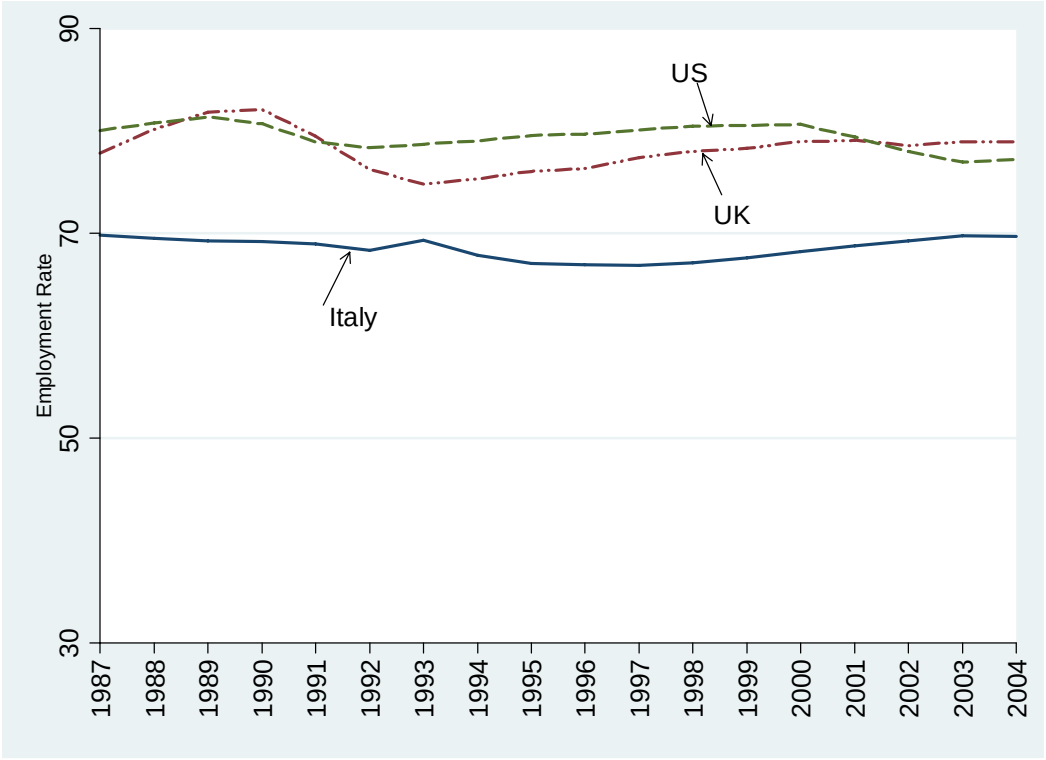
Notes: Based on CPS, GHS and SHIW. Samples include men in labour force, employed only, working full time, considering main job with positive potential experience.

Figure 2.2b: Percentile of Log Hourly Wage Distribution, Women



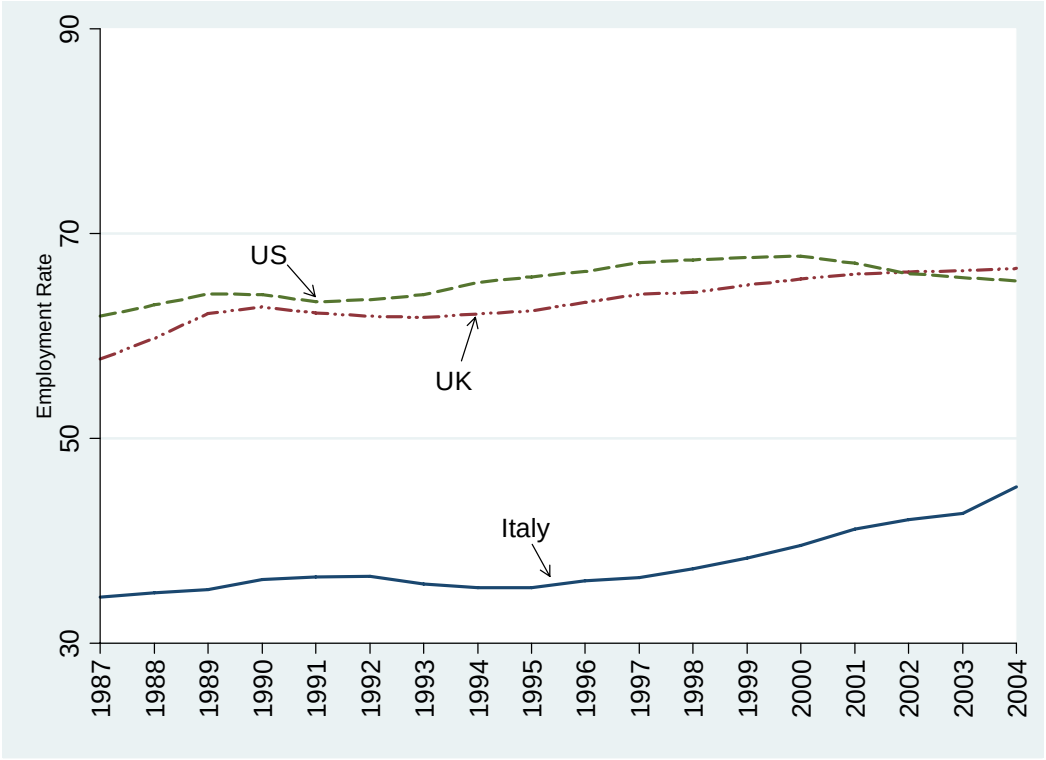
Notes: Based on CPS, GHS and SHIW. Samples include women in labour force, employed only, working full time, considering main job with positive potential experience.

Figure 2.3a: Employment Rate for Men, Italy, UK and U.S., 1987-2004



Notes: Based on the OECD data, employment refers to both part time and full time workers.

Figure 2.3b: **Employment Rate of Women, Italy, UK and U.S., 1987-2004**



Notes: Based on the OECD data, employment refers to both part time and full time workers

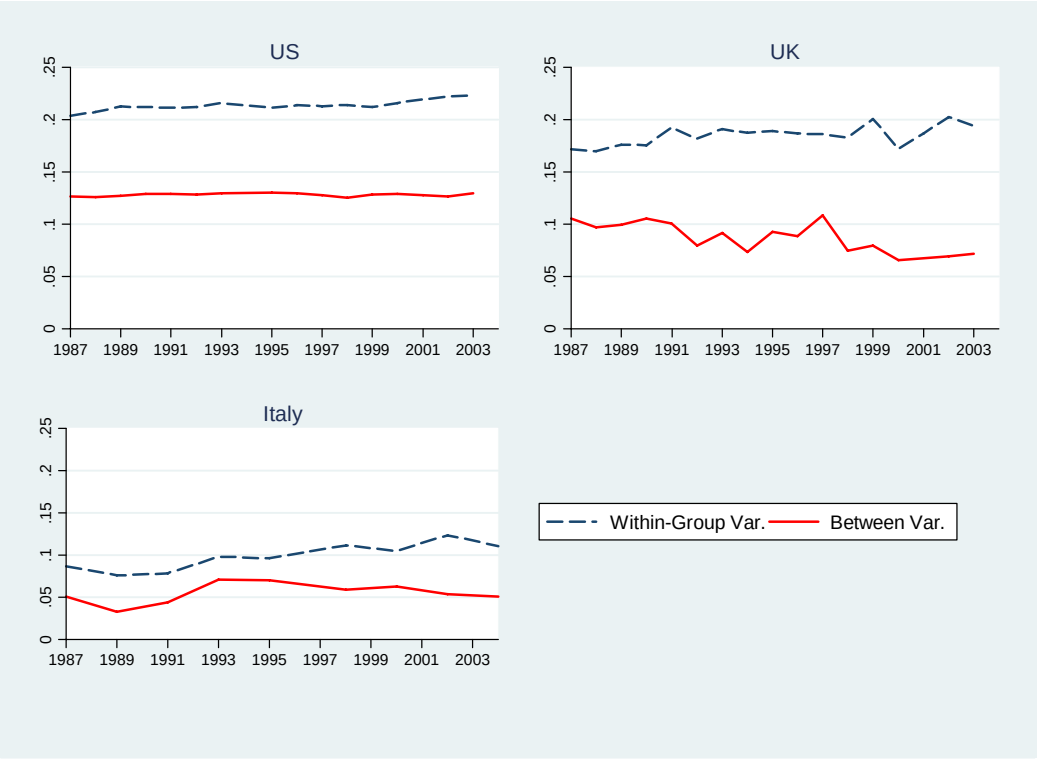
Table 2.1: Percentage Distribution of Workers by Education and Experience Groups

	Men		Women	
	1987	2003	1987	2003
	Panel A- US			
A. Education				
Lower	0.18	0.12	0.14	0.08
Intermediate	0.59	0.59	0.65	0.61
High	0.23	0.29	0.21	0.30
B. Experience				
1-10	0.36	0.28	0.38	0.28
11-20	0.30	0.26	0.28	0.24
21-30	0.18	0.26	0.19	0.27
31+	0.16	0.20	0.16	0.21
	Panel B- United Kingdom			
A. Education				
Lower	0.63	0.57	0.69	0.51
Intermediate	0.24	0.21	0.22	0.24
High	0.13	0.22	0.09	0.26
B. Experience				
1-10	0.30	0.20	0.47	0.25
11-20	0.23	0.24	0.18	0.25
21-30	0.21	0.28	0.17	0.26
31+	0.26	0.27	0.19	0.24
	Panel C-Italy			
A. Education				
Lower	0.58	0.45	0.41	0.28
Intermediate	0.31	0.44	0.42	0.54
High	0.11	0.11	0.17	0.18
B. Experience				
1-10	0.15	0.15	0.25	0.18
11-20	0.26	0.23	0.32	0.25
21-30	0.26	0.29	0.23	0.32
31+	0.33	0.32	0.20	0.26

Notes: The table shows the educational characteristics of full time workers, employed only. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder[26]. The higher education group refers either to individuals with graduate or postgraduate education.

[26] The statistics of this educational attainments comparison show that Italy stands out as a peculiar case with the level of higher education being lower than the UK and US and experiencing no changes over the time period analysed; moreover the table also presents a very low share of younger workers entering the labour market, reflected by the lower employment rate of the country. These surprising and disappointing figures are well documented by the OECD according to which the educational attainments in Italy are among the lowest in Europe and over the last few years these gaps have been growing. As pointed out by Naticchioni et al. (2008) this would suggest that the skilled workers are scarce in Italy.

Figure 2.4a: Within and Between-group Variance Men, 1987-2004



Notes: Based on a standard Mincer equation fit separately by year.

Figure 2.4b: Within and Between-group Variance Women, 1987-2004

Notes: Based on a standard Mincer equation fit separately by year.

Table 2.2a: R^2 of the OLS regressions

	United States		United Kingdom		Italy	
	Men	*Women*	*Men*	*Women*	*Men*	*Women*
Year				R^2		
1987	0.38	0.30	0.38	0.39	0.37	0.44
1988	0.38	0.29	0.36	0.34	-	-
1989	0.37	0.31	0.36	0.37	0.30	0.28
1990	0.38	0.31	0.38	0.38	-	-
1991	0.38	0.31	0.34	0.30	0.36	0.42
1992	0.38	0.31	0.30	0.33	-	-
1993	0.38	0.31	0.32	0.32	0.42	0.49
1994	-	-	0.28	0.34	-	-
1995	0.38	0.32	0.33	0.34	0.42	0.47
1996	0.38	0.33	0.32	0.35	-	-
1997	0.38	0.33	0.37	0.33	-	-
1998	0.37	0.33	0.29	0.28	0.35	0.34
1999	0.38	0.33	0.28	0.36	-	-
2000	0.37	0.33	0.28	0.24	0.37	0.39
2001	0.37	0.33	0.27	0.24	-	-
2002	0.36	0.33	0.26	0.26	0.30	0.32
2003	0.37	0.32	0.27	0.26		-
2004					0.31	0.33
Mean	***0.37***	***0.32***	***0.32***	***0.32***	***0.36***	***0.39***

Notes: Based on a standard Mincer equation fit separately by year.

Table 2.2b: R^2 of the OLS Regression for Pooled Years, UK

Year	**Men**	**Women**
1987	0.38	0.38
1988-1989	0.36	0.35
1990-1991	0.35	0.39
1992-1993	0.31	0.34
1994-1995	0.30	0.33
1996-1997	0.34	0.32
1998-1999	0.29	0.29
2000-2001	0.27	0.24
2002-2003	0.26	0.26
Mean	0.32	0.32

Notes: Based on a standard Mincer equation fit separately by year

Table 2.3a: Composition Effects and Changes in the Residual Variance of Log Hourly Wages, U.S.

	1987-1991	1991-1998	1998-2003	1987-2003
Men				
Actual Residual Variance	0.007	0.003	0.009	0.019 [86]
Residual Variance of 1987 skills distribution	0.003	-0.006	0.002	-0.002 [-0.01]
Residual Variance of 2003 skills distribution	0.007	-0.004	0.001	0.004 [0.18]
Total variance:	0.010	-0.001	0.013	0.022 [100]
Women				
Actual Residual Variance	0.007	0.002	0.009	0.019 [53]
Residual Variance of 1987 skills distribution	0.002	-0.010	0.002	-0.006 [-0.16]
Residual Variance of 2003 skills distribution	0.005	-0.003	0.002	0.005 [0.13]
Total variance:	0.013	0.010	0.011	0.035 [100]

Notes: Based on CPS. The table shows changes between 1987 and 2003 in wage decomposition. Residual variance is based on a standard Mincer equation. Numbers in squared brackets represent the percentage of the 1987-2003 change in the total variance of wages (both within and between-group components) that is attributable to this variance component.

Table 2.3b: Composition Effects and Changes in the Residual Variance of Log Hourly Wages, UK

	1987-1991	1991-1998	1998-2003	1987-2003
Men				
Actual Residual Variance	0.021	-0.009	0.011	0.023
Residual Variance of 1987 skills distribution	0.018	-0.019	0.005	0.003 [0.27]
Residual Variance of 2003 skills distribution	0.022	-0.019	0.014	0.016
Total variance:	0.015	-0.036	0.007	-0.011 [100]
Women				
Actual Residual Variance	0.013	-0.002	-0.012	-0.001 [20]
Residual Variance of 1987 skills distribution	0.010	-0.007	-0.013	-0.008 [17]
Residual Variance of 2003 skills distribution	0.007	0.001	-0.018	-0.010 [19]
Total variance:	-0.010	-0.012	-0.028	-0.053 [100]

Notes: Based on GHS. The table shows changes between 1987 and 2003 in wage decomposition. Residual variance is based on a standard Mincer equation. Numbers in squared brackets represent the percentage of the 1987-2003 change in the total variance of wages (both within and between-group components) that is attributable to this variance component.

Table 2.3c: Composition Effects and Changes in the Residual Variance of Log Hourly Wages, Italy

	1987-1991	1991-1998	1998-2004	1987-2004
Men				
Actual Residual Variance	-0.008	0.033	-0.001	0.024 [100]
Residual Variance of 1987 skills distribution	-0.008	0.030	-0.002	0.020 [83]
Residual Variance of 2003 skills distribution	-0.009	0.029	-0.001	0.019 [78]
Total variance:	-0.015	0.048	-0.009	0.024 [100]
Women				
Actual Residual Variance	-0.003	0.036	-0.014	0.019
Residual Variance of 1987 skills distribution	-0.005	0.039	-0.018	0.015
Residual Variance of 2003 skills distribution	-0.003	0.028	-0.013	0.013
Total variance:	-0.012	0.036	-0.022	0.003 [100]

Notes: Based on SHIW. The table shows changes between 1987 and 2004 in wage decomposition. Residual variance is based on a standard Mincer equation. Numbers in squared brackets represent the percentage of the 1987-2003 change in the total variance of wages (both within and between-group components) that is attributable to this variance component.

Table 2.4a: Within-group Variance of Wages by Experience-Education Cell for Men, U.S.

	Within-group variance			Work force share		
	1987-1989	2002-2003	Change	1987-1989	2002-2003	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Lower						
1-10	0.097	0.074	-0.023	0.062	0.045	-0.017
11-20	0.177	0.125	-0.053	0.034	0.023	-0.01
21-30	0.219	0.174	-0.044	0.026	0.023	-0.004
31+	0.233	0.185	-0.048	0.055	0.026	-0.028
Tot.	*0.726*	*0.558*	*-0.168*	*0.177*	*0.117*	*-0.059*
Intermediate						
1-10	0.166	0.143	-0.024	0.21	0.158	-0.051
11-20	0.191	0.196	0.005	0.181	0.148	-0.033
21-30	0.208	0.214	0.006	0.11	0.156	0.045
31+	0.242	0.247	0.006	0.085	0.126	0.042
Tot.	*0.807*	*0.8*	*-0.007*	*0.586*	*0.588*	*0.003*
High						
1-10	0.233	0.248	0.015	0.084	0.074	-0.01
11-20	0.254	0.297	0.043	0.087	0.089	0.002
21-30	0.299	0.322	0.022	0.043	0.082	0.039
31+	0.363	0.379	0.016	0.023	0.05	0.027
Tot.	*1.149*	*1.246*	*0.096*	*0.237*	*0.295*	*0.058*
				1	*1*	

Notes: Based on CPS, the table shows the within-group level and changes based on a standard Mincer equation by level of education and experience. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Table 2.4b: Within-group Variance of Wages by Experience-Education Cell for Women, U.S.

	Within-group variance			Work force share		
	1987-1989	2002-2003	Change	1987-1989	2002-2003	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Lower						
1-10	0.081	0.054	-0.027	0.049	0.037	-0.012
11-20	0.143	0.104	-0.039	0.023	0.013	-0.010
21-30	0.148	0.106	-0.042	0.024	0.015	-0.009
31+	0.153	0.14	-0.014	0.042	0.020	-0.022
Tot.	**0.525**	**0.404**	**-0.122**	**0.138**	**0.085**	**-0.053**
Intermediate						
1-10	0.152	0.135	-0.016	0.225	0.155	-0.070
11-20	0.201	0.185	-0.017	0.188	0.143	-0.045
21-30	0.196	0.2	0.004	0.134	0.172	0.038
31+	0.201	0.204	0.003	0.105	0.145	0.040
Tot.	**0.75**	**0.724**	**-0.026**	**0.652**	**0.615**	**-0.033**
High						
1-10	0.201	0.212	0.011	0.09	0.090	0.000
11-20	0.232	0.283	0.051	0.073	0.085	0.012
21-30	0.237	0.28	0.043	0.032	0.086	0.052
31+	0.259	0.271	0.012	0.015	0.039	0.024
Tot	**0.929**	**1.046**	**0.117**	**0.210**	**0.300**	**0.088**
				1	**1**	

Notes: Based on CPS, the table shows the within-group level and changes based on a standard Mincer equation by level of education and experience. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Table 2.4c: Within-group Variance of Wages by Experience-Education Cell for Men, UK

	Within-group variance			Work force share		
	1987-1989	2002-2003	Change	1987-1989	2002-2003	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Lower						
1-10	0.173	0.176	0.003	0.161	0.089	-0.072
11-20	0.156	0.171	0.015	0.114	0.134	0.02
21-30	0.185	0.181	-0.004	0.132	0.149	0.017
31+	0.161	0.191	0.03	0.205	0.209	0.005
Tot.	**0.675**	**0.719**	**0.044**	**0.612**	**0.581**	**-0.03**
Intermediate						
1-10	0.133	0.145	0.012	0.084	0.047	-0.037
11-20	0.123	0.232	0.109	0.082	0.059	-0.023
21-30	0.186	0.2	0.014	0.056	0.049	-0.005
31+	0.184	0.241	0.057	0.041	0.04	-0.001
Tot.	**0.626**	**0.818**	**0.192**	**0.263**	**0.195**	**-0.066**
High						
1-10	0.153	0.204	0.051	0.046	0.069	0.024
11-20	0.134	0.187	0.053	0.038	0.057	0.02
21-30	0.213	0.247	0.034	0.024	0.064	0.04
31+	0.23	0.222	-0.008	0.017	0.034	0.016
Tot.	**0.73**	**0.86**	**0.13**	**0.125**	**0.224**	**0.1**
				1	**1**	

Notes: Based on GHS, the table shows the within-group level and changes based on a standard Mincer equation by level of education and experience. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Table 2.4d: Within-group Variance of Wages by Experience-Education Cell for Women, UK

	Within-group variance			Work force share		
	1987-1989	2002-2003	Change	1987-1989	2002-2003	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Lower						
1-10	0.146	0.114	-0.032	0.287	0.087	-0.2
11-20	0.166	0.15	-0.016	0.109	0.104	-0.005
21-30	0.177	0.19	0.013	0.137	0.122	-0.016
31+	0.14	0.145	0.005	0.161	0.168	0.006
Tot.	**0.629**	**0.599**	**-0.03**	**0.694**	**0.481**	**-0.215**
Intermediate						
1-10	0.113	0.104	-0.009	0.113	0.077	-0.036
11-20	0.159	0.171	0.012	0.054	0.068	0.015
21-30	0.156	0.169	0.013	0.033	0.067	0.034
31+	0.195	0.154	-0.041	0.019	0.049	0.03
Tot.	**0.623**	**0.598**	**-0.025**	**0.219**	**0.261**	**0.043**
High						
1-10	0.133	0.097	-0.036	0.052	0.105	0.053
11-20	0.154	0.19	0.036	0.023	0.063	0.041
21-30	0.155	0.256	0.101	0.008	0.065	0.058
31+	0.221	0.176	-0.045	0.004	0.025	0.02
Tot.	**0.663**	**0.719**	**0.056**	**0.087**	**0.258**	**0.172**
				1	**1**	

Notes: Based on GHS, the table shows the within-group level and changes based on a standard Mincer equation by level of education and experience. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Table 2.4e: Within-group Variance of Wages by Experience-Education Cell for Men, Italy

	Within-group variance			Work force share		
	1987-1989	2002-2004	Change	1987-1989	2002-2004	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Lower						
1-10	0.082	0.191	0.108	0.055	0.033	-0.022
11-20	0.069	0.093	0.023	0.117	0.091	-0.027
21-30	0.07	0.105	0.035	0.137	0.133	-0.004
31+	0.077	0.108	0.031	0.253	0.203	-0.05
Tot.	**0.298**	**0.497**	**0.197**	**0.562**	**0.46**	**-0.103**
Intermediate						
1-10	0.075	0.087	0.013	0.09	0.099	0.009
11-20	0.074	0.096	0.022	0.103	0.116	0.013
21-30	0.108	0.12	0.012	0.09	0.129	0.039
31+	0.116	0.141	0.025	0.051	0.093	0.041
Tot.	**0.373**	**0.444**	**0.072**	**0.334**	**0.437**	**0.102**
High						
1-10	0.103	0.139	0.036	0.023	0.021	-0.002
11-20	0.106	0.111	0.005	0.039	0.027	-0.012
21-30	0.137	0.236	0.1	0.027	0.031	0.004
31+	0.149	0.167	0.019	0.015	0.024	0.009
Tot.	**0.495**	**0.653**	**0.16**	**0.104**	**0.103**	**-0.001**
				1	**1**	

Notes: Based on SHIW, the table shows the within-group level and changes based on a standard Mincer equation by level of education and experience. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Table 2.4f: Within-group Variance of Wages by Experience-Education Cell for Women, Italy

	Within-group variance			Work force share	
	1987-1989 (1)	2002-2004 (2)	Change (3)	1987-1989 (4)	2002-2004 (5)
A. By education and experience					
Lower					
1-10	0.087	0.128	0.041	0.076	0.021
11-20	0.070	0.089	0.019	0.089	0.052
21-30	0.068	0.088	0.020	0.100	0.078
31+	0.072	0.083	0.011	0.139	0.130
Tot.	**0.297**	**0.388**	**0.091**	**0.404**	**0.281**
Intermediate					
1-10	0.086	0.108	0.022	0.157	0.119
11-20	0.082	0.122	0.04	0.15	0.152
21-30	0.103	0.095	-0.008	0.087	0.178
31+	0.096	0.148	0.052	0.042	0.089
Tot.	**0.367**	**0.473**	**0.106**	**0.436**	**0.538**
High					
1-10	0.102	0.21	0.109	0.044	0.045
11-20	0.121	0.191	0.07	0.078	0.055
21-30	0.105	0.159	0.054	0.031	0.061
31+	0.109	0.179	0.07	0.007	0.020
Tot.	**0.437**	**0.739**	**0.303**	**0.16**	**0.181**
				1	1

Notes: Based on SHIW; the table shows the within-group level and changes based on a standard Mincer equation by level of education and experience. The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Figure 2.5a: Within-group Variance by Education, Men
(Average of the experience group)

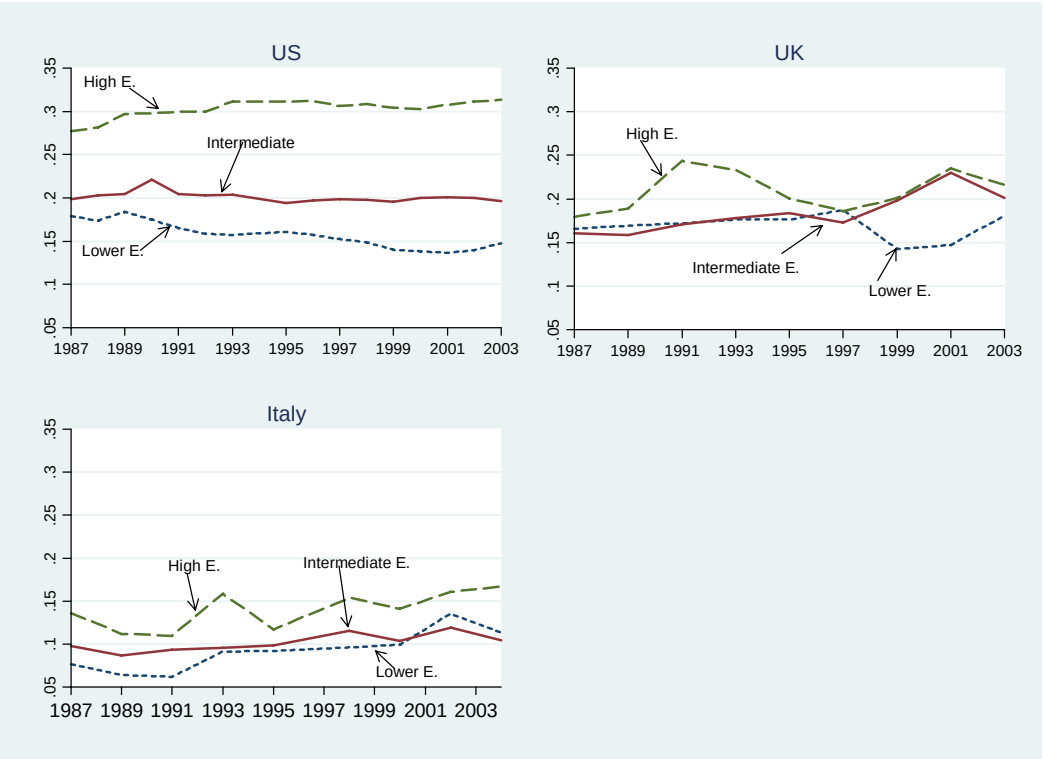
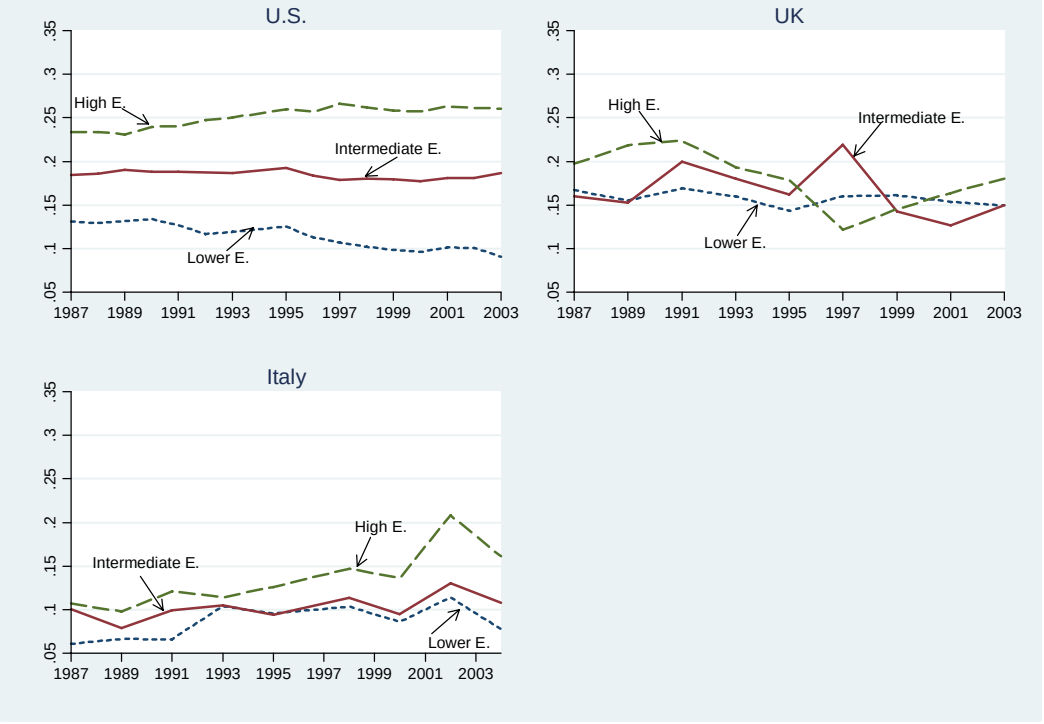
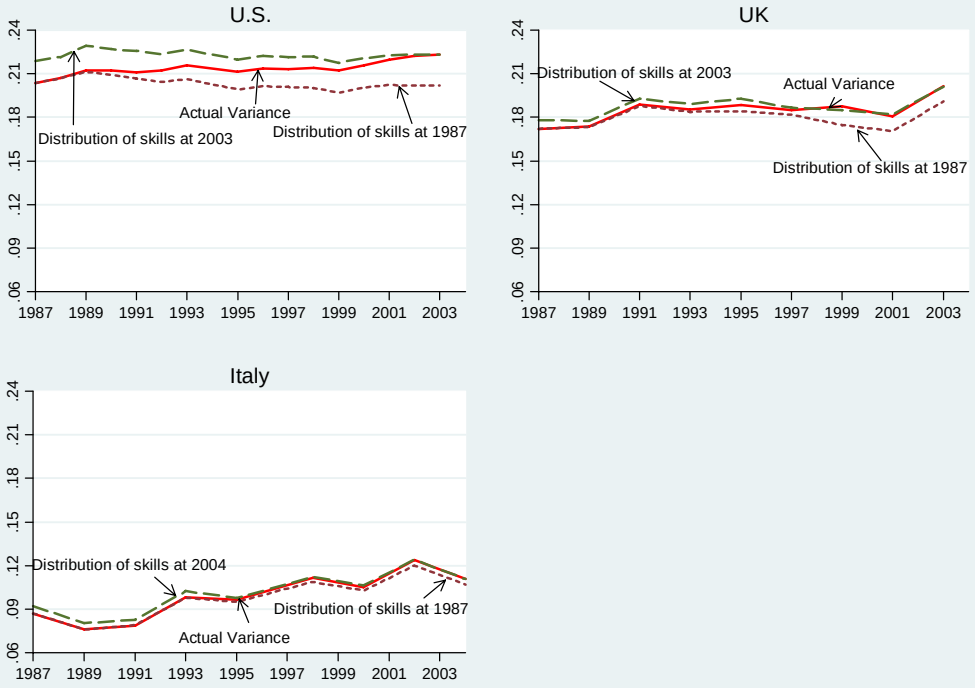


Figure 2.5b: Within-group Variance by Education, Women
(Average of the experience group)



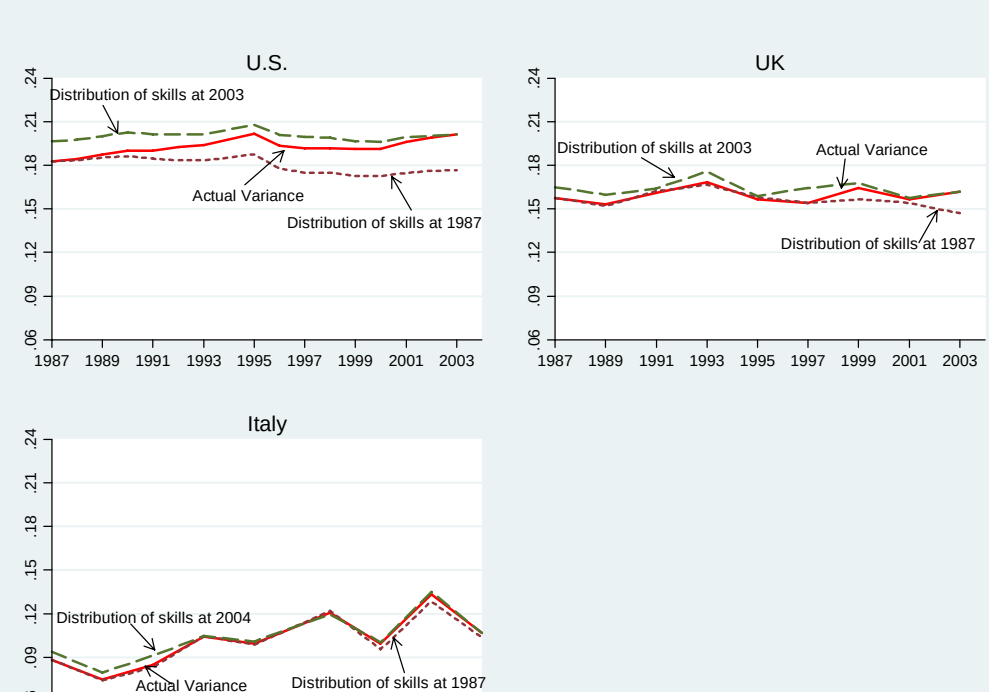
Notes: The table shows the within-group variance based on a standard Mincer equation; for each education group the within-group variance is the average of four experience groups (1-10, 11-20, 21-30, 31+). The lower education group includes workers who have completed compulsory education; the intermediate category includes workers with qualifications that exceed those of compulsory schooling but do not reach those of a college-degree holder. The higher education group refers either to individuals with graduate or postgraduate education.

Figure 2.6a: Actual and Counterfactual Residual Variance of Log Hourly Wages, Men



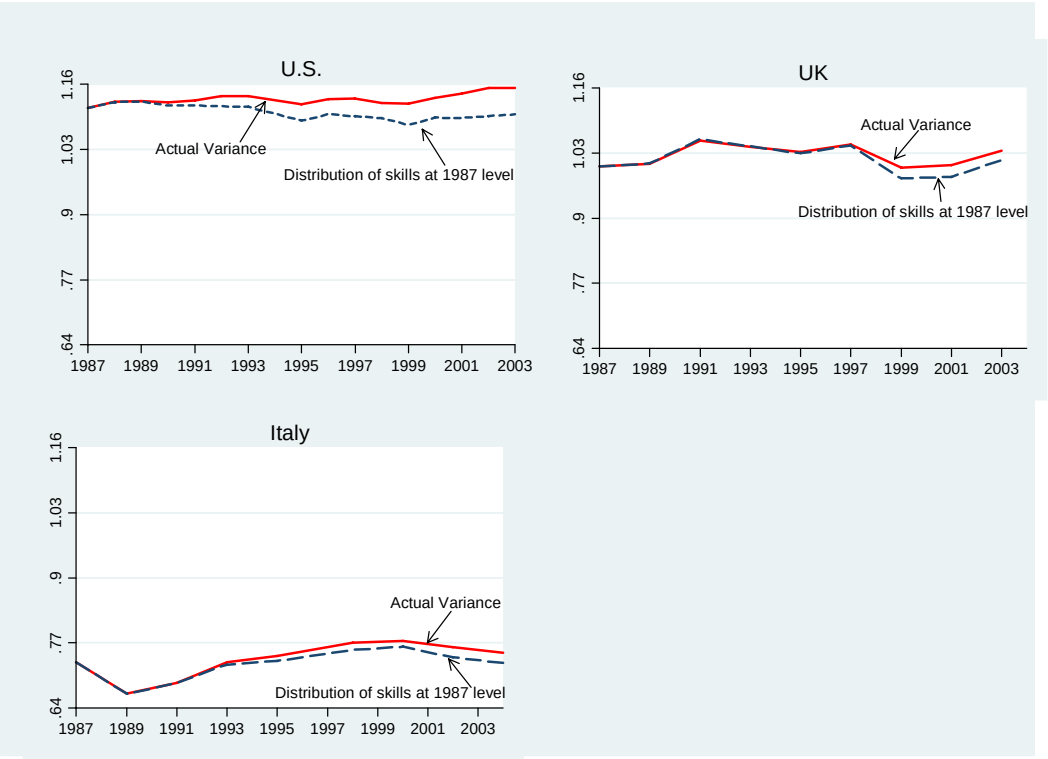
Notes: Based on CPS, GHS and SHIW. Samples include men of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.6b: Actual and Counterfactual Residual Variance of Log Hourly Wages, Women



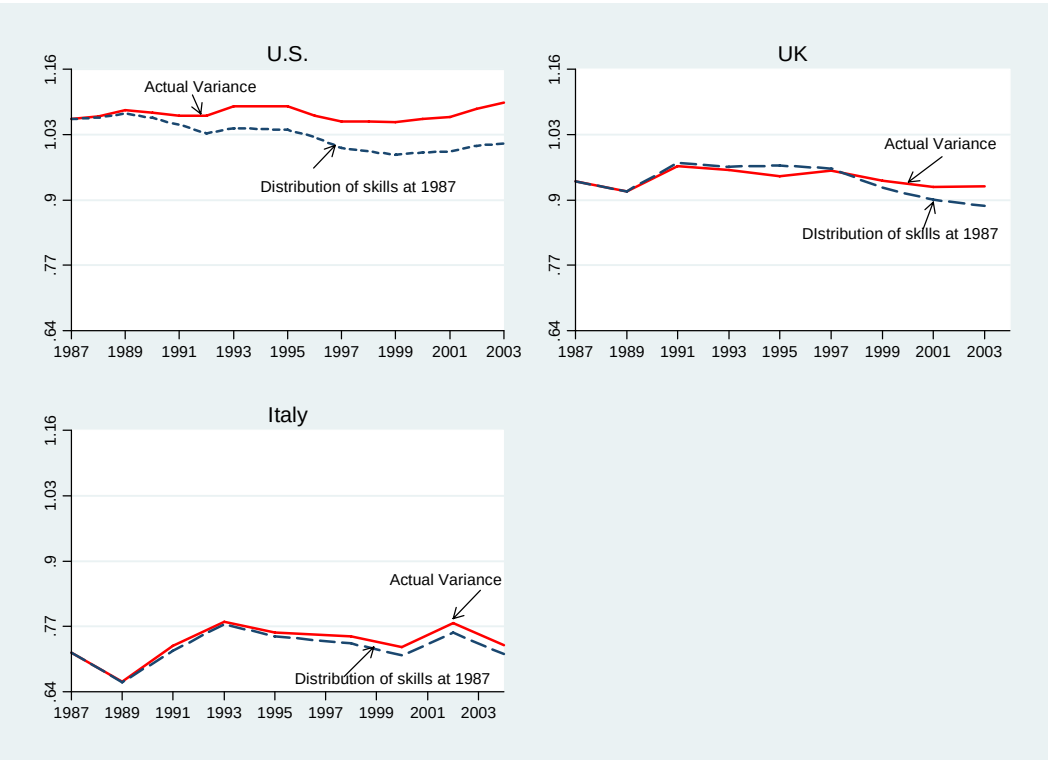
Notes: Based on CPS, GHS and SHIW. Samples include women of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.7a: 90-10 Residual Wage Gap, Men
(Holding distribution of skills at their 1987 level)



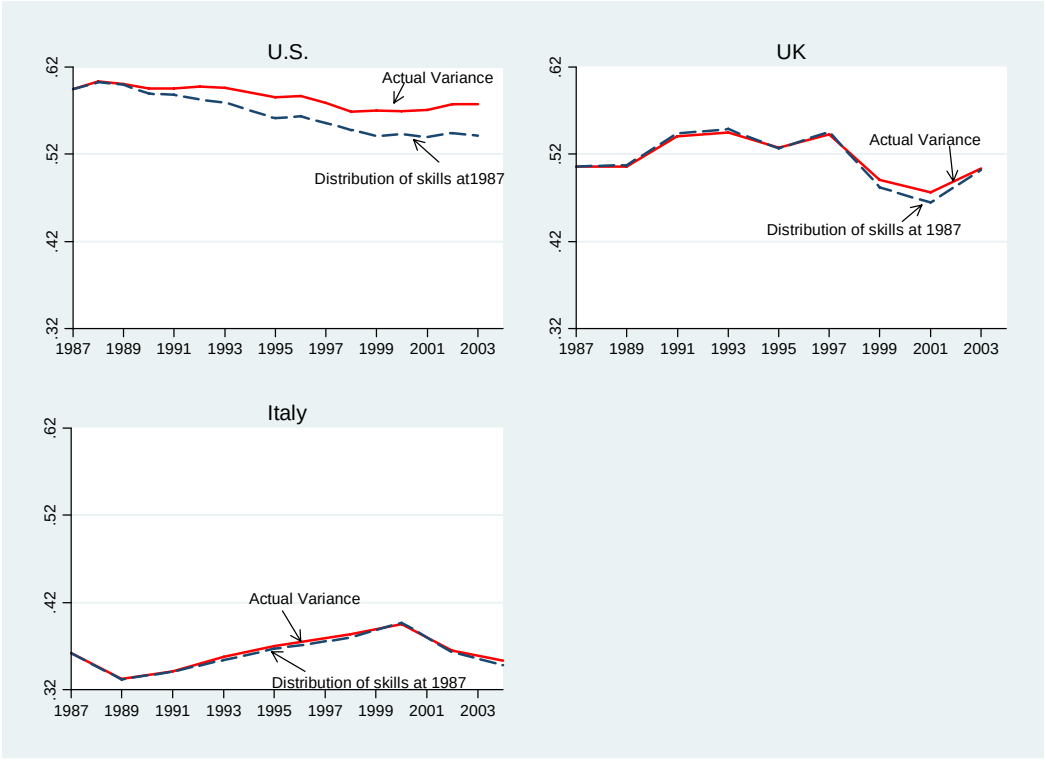
Notes: Based on CPS, GHS and SHIW. Samples include men of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.7b: 90-10 Residual Wage Gap, Women
(Holding distribution of skills at their 1987 level)



Notes: Based on CPS, GHS and SHIW. Samples include women of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.8a: 50-10 Residual Wage Gap, Men
(Holding distribution of skills at their 1987 level)

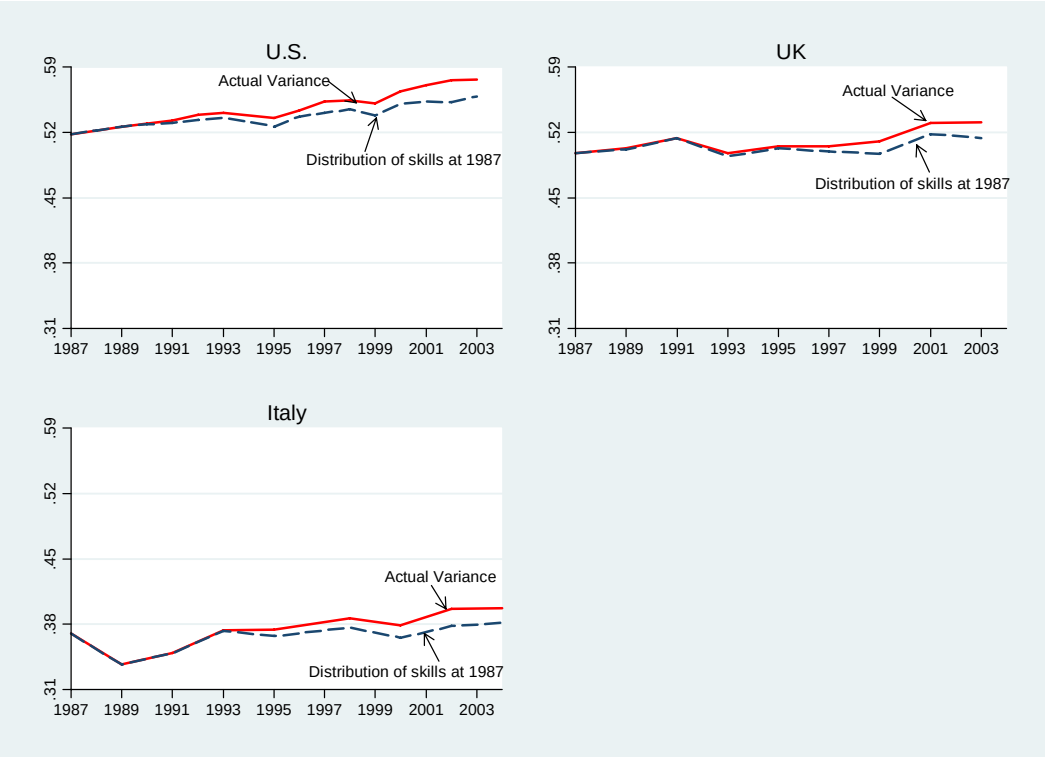


Notes: Based on CPS, GHS and SHIW. Samples include men of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.8b: 50-10 Residual Wage Gap, Women
(Holding distribution of skills at their 1987 level)

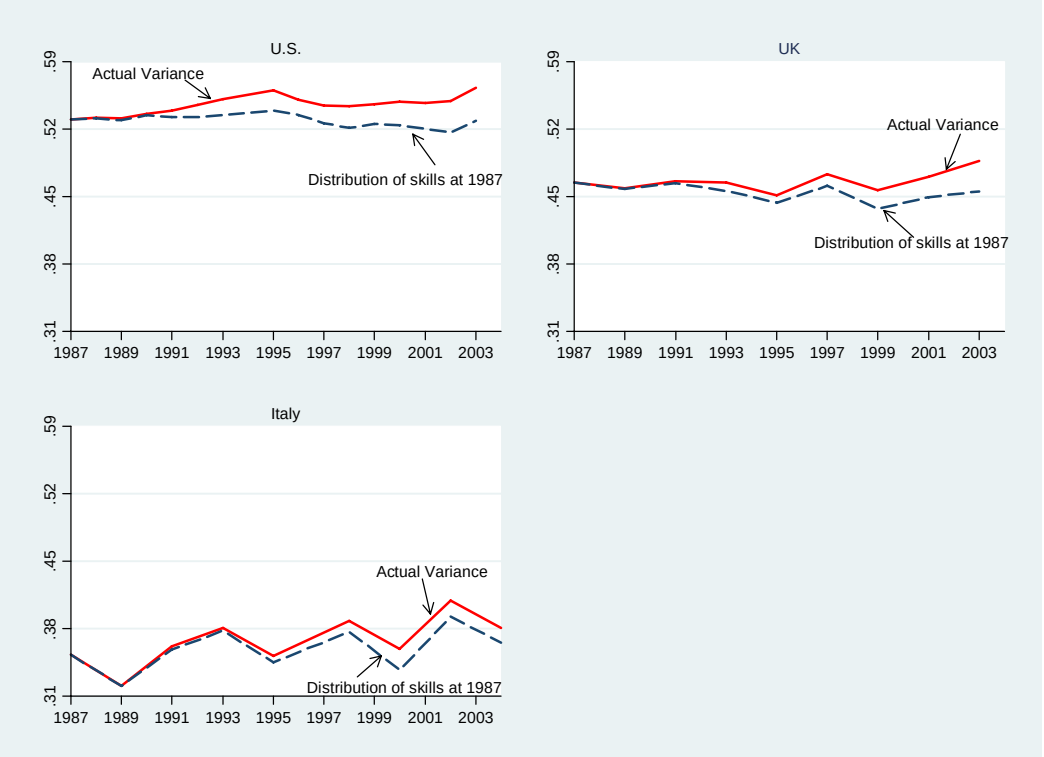
Notes: Based on CPS, GHS and SHIW. Samples include women of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.9a: 90-50 Residual Wage Gap, Men
(Holding distribution of skills at their 1987 level)



Notes: Based on CPS, GHS and SHIW. Samples include men of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Figure 2.9b: 90-50 Residual Wage Gap, Women
(Holding distribution of skills at their 1987 level)



Notes: Based on CPS, GHS and SHIW. Samples include women of labour force age, employed only, with positive potential experience, working full time and full year in their main job.

Chapter 3: Residual Wage Inequality and Immigration in the UK and the U.S.

3.1 Introduction

Over the last few decades, immigration has increased significantly in both the US and the UK; both countries have also experienced notable increases in the degree of wage inequality. According to Current Population Survey (CPS) and Labour Force (LFS) data, respectively, in 2008 nearly 14 percent of the US and 12 percent of the UK labour force[27] were born abroad. The same data also document that from 1994 to 2008, wage inequality, measured by the standard deviation of the natural logarithm of hourly wage for male workers, increased by 2.2 percentage points in the US and by 1.4 percentage points in the UK.

Overall, the increase in wage inequality has been stronger in the upper tail than the lower tail of the wage distribution (Machin and Van Reenen, 2008). Since the 1980s, both countries have experienced an increase in the upper-tail inequality (90-50), whereas the lower tail (50-10) has shrunk. Autor et al. (2008) document that for the US, between 1979 and 1987, the male 90/50 log hourly earnings rose by 8.5 log points; between 1987 and 1995, it rose by an additional 5.2; and it rose by 9.7 log points between 1995 and 2003. In contrast, between 1979 and 1987, the 50-10 gap increased by 13.0 log points, while the gap narrowed between 1987 and 2003 4.7 log points.

There is a huge empirical debate on the social and economic consequences of international migration, one of the core concerns relates to the impact of immigration on the wages of native workers. Despite the common-sense intuition behind the simple theoretical model of the laws of supply and demand, the international migration literature has found it difficult to arrive at a consensus on the impact of immigration on the wages of workers in the receiving countries.

Abdurrahman and Borjas (2007), for example, investigate the effect of the immigration-induced supply shifts on the cross-country evolution of relative wages earned in Canada, Mexico and the US. They find that the impact of migration on the wage structure differs significantly across the three countries. International migration reduced wage inequality in Canada, while it contributed to

[27] The sample includes full-time workers, who are not self-employed, have only a single job, have positive potential work experience, and receive an hourly wage less than 100 pounds for the UK and 100 dollars for the US

increased wage inequality in the US. International migration in Mexico increased the relative wages of workers in the middle of the skill distribution but lowered the wages of workers at the bottom of the distribution. Assuming that similarly educated workers with different experiences are not perfect substitutes, (but that immigrants and natives within each groups are), Borjas (2006) confirms the main predictions of the simplest theoretical model of a competitive labour market, reporting that in the US, a 10 percent increase in immigrant flow in the skill reduces weekly earnings by about 4 percent. In contrast, Card (2005) and Dustmann et al. (2005) give evidence that the effect of immigration on US and UK native wages is almost negligible.

A consensus emerging from a recent stream of the empirical literature is that immigration not only has little impact on natives' wages, but also, on average, exerts a positive rather than negative effect on natives' wages (Ottaviano and Peri 2006; Manacorda, Manning and Wadsworth 2007 (MMW hereafter); Dustmann, Frattini, Preston 2008 (DFP hereafter)). As proven by Ottaviano and Peri (2006) for the US and by MMW (2007) for the UK, natives' wage loss from immigration is mitigated by the incomplete substitutability of immigrants and natives within age and education groups. Ottaviano and Peri (2006) claim that imperfect substitutability may arise, among other aspects, from the different abilities or unobserved characteristics of workers. Recently, Borjas et al.(2008), re-examining the Ottaviano-Peri empirical exercise, show that their finding of imperfect substitutability is sensitive to the construction of the two key variables of relative wages and relative supplies—but particularly to the inclusion of young students in the sample. Under conventional classifications of workers by education and experience, the data fail to reject the hypothesis that immigrants and natives are perfect substitutes. Even allowing for long-run adjustments in the capital stock, immigration appears likely to lower the wages of those natives most affected by immigration-induced supply shifts.

The existing literature investigating the effect of immigration on native-born workers' wages has only addressed the role of observable characteristics (education and experience) of workers, while less attention has been devoted to

their unobservable skills[28]. Lemieux (2006) shows that residual (within-group) wage dispersion among workers with the same education and experience "…is generally believed to account for most of the growth in overall wage inequality". Consistent with the existing literature, Chapter 2 shows that the residual component accounts for most of the wage inequality variation in the U.S:, the UK and Italy. In the U.S. from 1987 to 2003, changes in the residual for men account for 86% of the overall change in the total variance, and 53% for woman. In the UK, despite the decrease in overall variance of wage both for men (0.033) and women (0.034), the residual variance increased by 0.014 of a per cent. More strikingly, the within-group variance in Italy explains almost all the change in wage inequality. During the past 20 years the UK and the U.S. experienced an increase in immigration. One may be tempted to think that the increase in immigrants over time in the U.S. and the UK might somehow be connected to the increase in residual wage inequality observed. Based on the Outgoing Rotation Group (ORG) for the US and the LFS data for the UK the correlation between changes in the residual and changes in the share of immigrants in the labour force between 1994 and 2008 is quite high; in the US, it is equal to 0.76 and 0.93, respectively, for men and women; in the UK, the correlation of changes is 0.88 for men and 0.96 for women, suggesting a possible positive relationship between the increased number of immigrants in the labour force and the increase in residual wage inequality.

Observed and unobserved distribution of skills are possibly larger for immigrants, therefore it is worth examining any link between residual wage inequality and immigration. This hypothesis is what this chapter attempts to test.

There is little existing research on the overall impact of immigration on native workers' within-group inequality. The only contribution has recently been provided by Card (2009), who shows that the relative level of residual wage inequality for natives in different skill groups is uncorrelated with the relative fraction of immigrants, suggesting that immigration has a relatively small causal effect even when studying the effect of immigration on within-group variation.

[28] Gould and Moav (2008), for example, empirically investigate how the emigration rate of Israeli workers in the US depends on their unobservable skills measured by residual wages from a standard Mincer equation.

Recently, Card (2009) offered an overview of the present understanding of the relationship between immigration and inequality, focusing on evidence from cross-city comparisons in the US.

Based on the fact that within-group wage inequality has risen substantially, he shows that across major cities, the level of residual wage inequality is strongly correlated with immigrants' local population share. In particular, a 10 percentage point increase in the immigrant share is associated with a 0.025-point increase in the residual variance of high school equivalent men's wages and a 0.027-point increase in the residual variance of college equivalent men's wages. However, the effects of immigration on within-group inequality are small: immigrants account only for 5% of the increase in US wage inequality between 1980 and 2000, although immigrants tend to have higher residual inequality than natives.

There are many reasons to believe that immigrants and natives with similar observable skills may report differences in wages because of differences in unobservable skills that are relevant in the labour market. Immigrants constitute a particular subgroup of their original population with motivations and tastes that may distinguish them from natives (Ottaviano and Peri, 2006). In manual and intellectual work, they have culture-specific skills as well as limits (language) that might translate into advantages or disadvantages (Peri and Sparber 2008); foreign-born workers have different abilities pertaining to language, quantitative skills, and relational skills, so they choose occupations differently from natives, even within the same education and experience group.

Using LFS data, Clark and Lindley (2006) document the impact of economic conditions of year of arrival and assimilation on the labour market outcomes of immigrants to the UK from 1993 to 2002. They find positive earnings assimilation for all immigrant groups and strong employment assimilation for those who completed their education in the UK. They explain that after arrival in the host country, immigrant labour market outcomes will adjust toward those of non-immigrants or native workers. Assimilation is thought to take place through human capital enhancement: immigrants acquire skills that are specific to the destination country, including knowledge of the labour market and language proficiency. Preliminary findings by Chiswick (1978) for the US found

that while immigrants earned significantly less than natives upon arrival, they caught up with natives in terms of earnings as they integrated in the host country.

Another key fact that contributed to the debate on both wage inequality and immigration is the changes in the characteristics of workers. The labour force has been growing older and more educated (Autor et al., 2005; Lemieux 2006, 2008); Lemieux (2006) argues that these secular changes in the education and age structure may mechanically increase residual wage inequality; he shows that a large fraction of the 1973-2003 growth in the residual wage inequality in the US is a "spurious" consequence of composition effects. The methodology requires taking the actual residual variance of the log hourly wage OLS regressions and re-weight it, holding the characteristics of the labour force constant at a base year.

However, this influential contribution does not account for the fact that in recent decades, not only has the share of immigrants in the labour force been increasing, but there have also been changes in the characteristics (education and experience) of immigrants. An extensive literature (Dustmann, Fabbri, Preston 2005**;** DFP 2008; MMW 2007; Schmitt and Wadsworth 2007; Wadsworth 2010) shows that compared to people born in the UK, immigrants are, on average, better educated. Similarly, for Canada, Boudarbat and Lemieux (2008) explain that in terms of years of completed education, immigrants are more educated than the Canadian born, and the education gap is growing with time. Immigrants are also typically younger than the native-born workforce, on average.

This chapter aims to adapt one of the main challenges of the 1990s wage literature to the immigration context: wage dispersion is not fully explained by variables linked to the standard human capital model, like education and experience. The residual or within-group wage inequality[29] – wage dispersion among workers with the same education and experience - accounts for most of the growth in overall wage inequality (Juhn et al. 1993; Acemoglu 2002; Autor et al. 2005, 2008; Lemieux 2006).

Unlike previous studies, this paper focuses on the effects of immigration on residual wage inequality in the UK and US between 1994 and 2008. It seeks to assess whether and to what degree immigration contributed, along with

[29] Intuitively, years of schooling and experience do not capture returns to other skills. In contrast, regression-based residuals include unmeasured aspects of human capital such as school quality, ability, effort and innate skills.

technology, institutions and traditional explanations, to widening inequality. To do so, this work reassesses Lemieux's hypothesis (i.e. that composition effects exert an upward mechanical force on the residual wage inequality) by adding the immigration dimension to the original analysis. This work differs from that of Card (2009) in different ways: first, it differs in the methodology adopted; second it differs in the sample period analysed in fact while Card considers 1980, 1990, 2000 Census and 2005/2006 American Community Survey this work proposes a time series analysis from 1994 to 2008; finally it also presents a comparison with the UK that similarly to the US experienced high increase in wage inequality and immigration.

The methodology used by Lemieux (2006) and implemented in this and the previous chapter is based on a simple approach that can analyse changes in the distribution of wages that are economically interpretable using the standard tools of human capital theory. Unlike other methodologies (Juhn et al. 1993; Autor et al.2005- 2008; Melly 2005), the procedure controls for changes in the distributions of observables and the share of immigrants in the labour force by holding the skill distribution of the workforce and the supply of foreign-born workers constant at a base year; this method requires re-weighting the actual residual variance of the log hourly wage OLS regressions according to a weight that holds the characteristics of the labour force constant at a base year. In addition to the original methodology and extending the work of Chapter 2, this chapter controls for an increasing supply of immigrants by constructing another weight that fixes the share of immigrants at a base year.

The chapter also illustrates the evolution of the upper- and lower-tail distributions when controlling for both composition effects and the increasing supply of immigrants. To account for the fact that immigrants perform differently[30] according to the amount of time they have already spent in the receiving country (DFP 2008; Schmitt and Wadsworth 2007) and to account for any cohort effects, the analysis is run separately for recent and other immigrants,

[30]DFP (2008) demonstrate that immigrants downgrade considerably within educational categories upon arrival, particularly because of their lack of complementary skills like language. Recent immigrants may not be able to make use of their educational background to its full potential.

defined, respectively, as those who have spent five years or less in the US or UK and those who have been in the US or UK for more than five years[31].

The structure of this Chapter is as follows. The next section provides an explanation of why unobservable components might be crucial in analysing the impact of immigrants on natives' wage inequality and what the plausible link is between the increased educational attainment of the labour force and the increase of the residual; part three presents the econometric methodology; part four describes the two datasets used in this paper; part five discusses the results; and the final section concludes.

3.2 Immigrants and Natives: Imperfect Substitutability, Unobserved Skills and Composition Effect

The textbook model of a competitive labour market has unambiguous implications for migration-induced supply shift effects on natives' wages and employment opportunities. Labour inflow should lower the wages of competing workers. However, a key determinant of the positive or negative effect of immigration inflow on natives' wages is the degree of substitution between immigrant and native workers. Holding capital constant and assuming constant returns to scale production technology, an increase in the labour supply due to an immigrant inflow will lower wages if immigrants and natives are substitutes. In contrast, if immigrants and natives are imperfect substitutes in production, then the increase in the labour supply due to immigration can boost natives' wages (Borjas 2008).

The lack of any negative effect of immigration on wages of natives in the US and the UK reported by Ottaviano and Peri (2006) and MMW (2007) is based on the evidence that immigrants do not fully compete and substitute with natives, even within a given education-experience group.

[31] The existing literature defines earlier and new immigrants in different ways; DFP (2008) define earlier and recent immigrants in the UK, respectively, as those who have been in the UK two years or more and those who arrived within the last two years. Chiquiar and Hanson (2006) define recent Mexican immigrants in the US as those who migrated in the last ten years. Schmitt and Wadsworth (2007) consider recent immigrants as those who have spent up to five years in the US or UK.

As highlighted by Ottaviano and Peri (2006), the imperfect substitution between immigrants and natives may be due, among other reasons, to their differences in unobservable skills that are relevant in the labour market. Immigrants are a particular group of their native country's original population with motivation and tastes that may differentiate them from natives of the host country. In manual and intellectual work, their culture-specific skills and limitations (e.g., language) might translate into advantages or disadvantages; foreign-born workers have different abilities pertaining to language, quantitative skills, and relational skills, so they choose occupations differently from natives, even within the same education and experience group.

Not dissimilar to the spirit of Ottaviano and Peri (2006); Gould and Moav[32] (2008) argue that unobservable skills are a mixture of "general" skills (like education) that can be easily transported to another country and other "country-specific" skills that cannot be easily transported to another country; "country-specific" skills include personal connections, knowledge of the local labour market, language-specific communication skills, and success in the labour market. A significant proportion of an individual's total human capital is likely to be country specific for several reasons; language and cultural barriers may prevent an individual from transferring his or her skills to a country where he/she lacks a commanding knowledge of the local languages, consumer tastes and so on. The idea that there might be a crucial unobservable skills component in the immigration process is consistent with a recent consensus of the wage inequality literature: unobservable skills, measured by the residual of a Mincer equation, explain most of the variation in the observed increasing wage inequality (Acemoglu 2002; Juhn et al.1993; Autor and Katz 1999; Autor et al., 2005, 2008;Lemieux 2006).

Evidence from the US (Autor and Katz 1999) shows that residual wage inequality started increasing in the 1970s and continued rising considerably in the 1980s and then at a slower pace in the 1990s; the residual log weekly wage inequality for full-time, full-year workers increased by 27 log points for men and

[32] They show that a higher ratio of individuals with a higher transferability rate of unobservable skills exists in the middle of the distribution of total unobservable skills. They argue that those at the bottom of the unobservable skills distribution have little of both types of skills, while those at the top have high levels of both. Individuals are more likely to have high levels of unobservable general skills versus country-specific skills if they are in the middle of the distribution rather than on the tails.

25 log points for women from 1963 to 1995. For the US, Juhn et al. (1993) show that unobserved components affect the wages of workers at the top and at the bottom of the distribution to different extents. Over the period of 1964-1988, changes in unobservable quantities accounted for 65 percent of the increase in inequality for workers below the median but for less than half of the increase in inequality for those above the median.

More recently, Lemieux (2006) outlines the main patterns observed in the United States using CPS data: the residual variance for full-time, full-year male workers from 1973 to 2003 grew by about 0.04 log points; most of that growth was concentrated in the 1980s, remained essentially unchanged during the 1990s and grew again between 1999 and 2003 Chapter 2 of this thesis confirms this basic finding for the UK, the US and Italy.

The leading explanations of wage inequality, such as declining unionisation, the falling real minimum wage and the SBTC[33], thus do not seem to help explain the main recent trends in wage inequality because they do not readily predict increasing upper-tail inequality and compressing lower tail inequality (Autor et al.2005). As documented by Lemieux, there are two main factors affecting the increase in residual wage inequality over the last few decades:

i) The "price" or return to unobserved skills may be increasing over time because of the increase in the demand for skills;

ii) Dispersion in the residual could be increasing because of composition effects.

Changes in characteristics[34] affect both the demand and supply of observed and unobserved skills and can alter wage and employment outcomes (Autor and Katz, 1999). Changes in within-group inequality may reflect market forces changing the returns to (unmeasured) skills. Therefore, the increase in within-group inequality can be interpreted as reflecting an increase in the returns

[33] Similarly, Lemieux (2006) discusses why the SBTC explanation presents several limitations; first, it could not help to explain some diverging patterns of inequality across advanced countries; SBTC depended on the weak fact that the residual inequality had been increasing since the 1970s. The failure of the SBTC for developing countries can be explained by the fact that because those countries are subject to the same technological change, they did not experience an increase in inequality as we should expect. They did not vary in a similar way over time.

[34] Acemoglu (2002) explains that composition effects cannot by themselves explain the recent changes in wage dispersion, but suggests that inequality among more educated and less educated workers should move in opposite directions.

to unobserved skills. Holding market prices constant, changes in labour force composition can mechanically raise or lower overall earnings dispersion by increasing or reducing heterogeneity in observed skills (education and experience).

The link between composition effects and the residual can be explained by the fact that when the level of education of the labour force increases, more and more "marginal" workers are added to the group of highly educated workers, creating more unobserved heterogeneity in that group and increasing within-group inequality. Therefore, an increase in the supply of more educated workers will immediately benefit those workers with more unobserved skills and will also depress returns to schooling while increasing within-group inequality. The more skilled workers within each education group also benefit from skill-based technical progress. Technical change spurred by the increase in the supply of educated workers will immediately benefit workers with more unobserved skills, raising within-group inequality. Therefore, an increase in the supply of educated workers will depress the returns to schooling while increasing within-group inequality[35].

An increase in the proportion of the workers with more education and experience can also mechanically raise residual wage inequality because earnings variation is higher for college-educated relative to high school-educated workers (Machin and Van Reenen 2008).

A number of influential studies (see, for example, Juhn et al.(1993) Autor et al.(2005)) focus on this interplay between changes in educational and experience characteristics of the labour force and the evolution of the residual to evaluate how much of the overall increase in wage inequality can be attributed to wage dispersion among workers with the same education and experience and how much is due to changes in the composition of the workforce. The evidence provided is sometimes contradictory, presumably due to the different methodologies applied.

[35] Mincer (1998) applies a human capital analysis to intra-group wage inequality, measured by variances in log wages, and their changes over time reveal the US wage structure changes from 1970-1990. In a similar vein to Acemoglu, he provides evidence that within-group inequality is not directly or closely related to between-group variances; therefore, we can expect differences in the fluctuations of the two components of inequality.

Lemieux (2006) concludes that a large fraction of the 1973-2003 growth in residual wage inequality in the US is due to composition effects, affecting both the upper and lower tails of the distribution. In particular, he demonstrates that the increase in within-wage inequality appears to be a spurious consequence of the fact that workforces became older and more educated over time; in other words, the increases in educational level and experience act as a mechanical force on the residual such that the increase in the variance of the residual carries the same sign as the change in the composition of the workforce. This phenomenon means that changes in the education and experience characteristics of the workforce determine more variation in wages due to unmeasured aspects of human capital.

There is no theoretical model explaining how the increasing supply of immigration could affect changes in the residual wage inequality of the labour force. However, because immigrants and natives differ in their unobservable skills if they are imperfect substitutes within education and experience group, we might expect immigrants to have more heterogeneous unobservable skills than natives even within the same education group; when the share of immigrants in the labour force increases, we could expect this to create even more variation in the overall residual wage inequality.

3.3 Econometric methodology

This chapter seeks to test Lemieux's findings discussed in chapter 2 by adding the immigration context specifically it aims to measure how much of the increase in the residual wage inequality observed in the UK and US between 1994 and 2008 is due to the increasing share of immigrants in the labour force; in similar vein to chapter 2, the econometric methodology of this chapter is based on two simple steps: the analysis of inequality in the residuals and a reweighting approach to control for both composition effects and increasing supply of immigration in both the US and the UK.

Because the analysis of this chapter adds the immigration dimension, the methodology adopted displays some differences with the one adopted in the previous chapter.

First of all the variance decomposition is carried out separately for the whole labour force (natives and immigrants together); for natives only, for immigrants

only, and by gender. The analysis of inequality in the residual by immigration status depicts the differences in the levels and trends of wage inequality experienced by natives and immigrants in both the UK and the US, moreover it also allows to measure how much of the observed wage inequality in the whole labour force is due to immigrants. The second substantial difference between the methodology adopted in this chapter and the one adopted in chapter 2, is the extension of the reweighting approach to the immigration status, in other words the reweighting approach is used to control for the increasing supply of immigrants helping therefore to disentangle the effect of immigrants on the increasing residual wage inequality.

Defining δ_m as the fraction of immigrants in the labour force, the residual variance defined in (6) of Chapter 2 can be re-written as:

(1) $\quad Var(\varepsilon_{it}) = \Sigma\delta_{mt} V_{jt}$

where the residual variance is now a function of the variances of wages accounting not only for the skill characteristics of the labour force but also for the share of immigrants in the labour force.

In a similar vein to (9) in Chapter 2, by holding fixed the share of immigrants in the labour force (δ^*_{mt}), the counterfactual residual variance can be written as:

(2) $\quad V_t^* = \Sigma\, \delta^*_{mt} V_{jt}$

Working in this way will investigate whether the increasing share of immigrants in the labour force can help to account for the increasing residual wage inequality observed.

Controlling for the increasing supply of immigration can be done by following the same rewighting methodology followed in chapter 2,but now constructing another weight that, in the spirit of DiNardo et al. (1996), in addition to the usual education-experience variables adds a dummy variable for the state of immigration as well as some interaction terms between education and the area of origin of immigrants to make the logit flexible enough. The magnitude of the weight therefore depends on the share of immigrants in the labour force observed in year *t* compared to the share of immigrants observed in a given base year *s*.

Working in this way would allow the construction of the immigration counterfactual: "what would have happened to the residual wage inequality of the labour force had the share of immigrants remained constant at a base year".

3.4 Data

The analysis is based on two pooled time series cross-sectional micro data sets: the Outgoing Rotation Group (ORG CPS) for the US and the Labour Force Survey (LFS) for the UK. Both data sets cover the same period (1994 to 2008), contain similar information relevant to labour market characteristics, wages and immigration, and are sufficiently large to analyse minority populations. The analysis starts from 1994 because the main dataset for the U.S. (ORG CPS) do not contain information on the immigration status of individual in the previous years The CPS is a monthly household survey conducted by the Bureau of Labour Statistics to measure labour force participation and employment. The survey provides individual data for about 30,000 individuals per month. Every household that enters the CPS is interviewed each month for four months, then ignored for eight months, then interviewed again for four more months. Standard weekly hours/earnings questions are asked only during a household's fourth and eighth interviews. These outgoing interviews are the only ones included in the extracts. New households enter each month, so one fourth of the households are in an outgoing rotation each month.

One of the main advantages of the ORG CPS is that it provides point-in-time measures of usual hourly wage for 60 percent of the sample; the remaining non-hourly wage can easily be calculated as the ratio of earnings to hours. We kept real hourly wage between $1 and $100. One of the main issues encountered when working with the ORG CPS dataset relates to the top-coded earnings. In the data released to the public, the Census Bureau restricts the top of the earnings distribution to $99,999 a year. This restriction means that all earnings above that level appear in the CPS public dataset as $99,999, whatever the actual earnings are. This artificial ceiling can lead to bias in the measurement of trends in earnings inequality if the proportion of earnings so affected changes over time; in particular, it will lower the mean and the variance of the wage data relative to the true mean and variance. We adjust for the top-coding issue by using the log-

normal approach recommended by Schmitt (2003). In contrast[36] to the procedure that is usually applied, the log-normal procedure models the entire distribution, not just the top portion of interest, under the assumption that the entire distribution of earnings is log-normally distributed. The properties of the log-normal distribution allow for the straightforward estimation of the mean and variance of the "true" distribution, even though the estimates of the mean above the top-code are consistently below those generated by these versions of the pareto approach.[37]

The LFS is very similar to the US Current Population Survey in terms of its purpose (measuring labour market activity and unemployment in a timely fashion), sample size and because, similar to CPS, LFS provides point-in-time hourly wages for a large fraction (around 40%) of each sample.

The (LFS) is the largest survey of households living at private addresses and in NHS accommodations in the UK and is conducted by the Office for National Statistics (ONS). Information is recorded in four quarters; each quarter's LFS sample of 53,000 UK households consists of five "waves", each of approximately 11,000 private households. Each wave is interviewed in five successive quarters, and 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 postcode. To limit the effects of outliers, following the existing literature in the UK (MMW 2007), only observations with an hourly wage between 1 and 100 pounds in 2008 pounds are kept. In the same manner as for the CPS, for individuals whose wage is only recorded weekly, hourly wage is derived by dividing weekly wage by the usual amount of paid hours worked per week.

Real wages for the UK are obtained by deflating nominal wages by the Retail Price Index. For the sake of comparability with the UK, wages measures for the US are deflated using the Consumer Price Index.

The samples used for the estimations consist of men and women in the labour force, considered separately, meaning men and women aged 16-64 for the

[36] A large part of the existing literature on wage inequality (Lemieux 2006, Katz and Autor 1999, Autor et al.2005,2008) addresses the top-coding issue by multiplying top-coded wages by a factor of 1.3 or 1.4, which is believed to provide estimates of the mean and the variance that are closer to their true values.

[37] For details, see John Schmitt (2003): "Creating a consistent hourly wage series from the Current Population Survey's Outgoing Rotation Group, 1979-2002".

US and aged 16-64 and 16-59 respectively, for the UK. We limit the analysis to workers who are full-time employees,[38] considering only their main job. We define as an immigrant someone who was born outside of the US and UK irrespective of the time of age on arrival.

All results are derived from separate regressions for men and women, respectively, for all workers, natives, and immigrants by applying the log hourly wage to a set of dummies for age, education and interactions between education and age squared.

The similarity between the US and UK schooling systems allowed us to create three educational categories that are broadly comparable. The lower education group includes workers who have completed compulsory education, i.e., less than a lower secondary education; for both the US and the UK, this group corresponds to 0 to 11 years of schooling. The intermediate category includes workers with qualifications that exceed those of a high-school dropout but do not reach those of a college-degree holder (both excluded). In both countries analysed, this corresponds to any individual with years of schooling at least 12 and at most 15 years of schooling. The high education group refers either to individuals with graduate or postgraduate education, corresponding to 16 or more years of schooling. One of the main issues arising when aggregating immigrants and natives based on level of education is that, due to the heterogeneity of educational systems, there is not a one-to-one correspondence in years of schooling. In the ORG/CPS data, education is reported in years for all workers, meaning that immigrants report their level of education in terms of years of schooling. One problem arising for LFS is that foreign educational qualifications are classified in the "other" category. As explained by MMW (2007), there is good reason to believe that many immigrants in the "other" category actually have quite high levels of education. The LFS employs an alternative definition of educational level, namely, the age at which the individual left full-time education. To create comparable educational categories for the UK data, we combined information on "age left school" and "other". In particular for immigrants reporting the "other "category of education, we defined immigrants to have higher level of education if they left school at age 21 or more; immigrants to have intermediate level of

[38] The use of only full-time workers is meant to eliminate variation in hours worked per week or weeks worked per year (Katz and Murphy, 1992; Card and Lemieux 2001).

education if they left school between the age of 17 and 20 and finally we define immigrants to have lower level of education if they left school before the age of 16.

The years of potential labour market experience variable is conventionally derived as age – years of completed education- the age at which children start school. Workers are aggregated into four-year experience intervals (0-10; 11-20; 21-30; 31+). Based on the three education categories (lower, intermediate and high) and the four experience categories, workers can be classified into one of twelve skill groups[39].

3.5 Results

3.5.1 Characteristics of Immigrants and Natives

Tables 3.1 to 3.3 outline the composition of and changes in the workforce over time in the UK and U.S. As documented in Table 3.1, during the period analysed in both countries, the number of immigrants increased; in the UK, the share of immigrants in the labour force more than doubled between 1994 and 2008; for example, the presence of immigrants in the UK male labour force increased from 6% in 1994 to 12% in 2008; similarly, for females, the share of immigrants increased by 7 percentage points. In the US, the corresponding increases in labour force share were from 10% to 15% for male workers and 4% for females. One of the reasons for the resurgence of interest in immigration, in addition to its volume, is the composition of the immigrant population. Table 3.2 documents changes in the areas of origin of immigrants, demonstrating that rising immigration is also associated with a change in the mix of immigrants' origins. In the US, the trend is for more immigrants from the Americas at the expense of Europe; for the UK, the reverse is true, with almost half of immigrants coming from Europe while the share of Americans decreased over the sample period. The increase of Europeans in the UK in 2008 largely reflects the arrival of A8 accession countries that were given free movement of labour after 2004.

[39] In the original data, Lemieux defines 20 education and experience groups based on 5 education groups and 4 experience groups. The education groups are dropout; high school graduates; ; some college; college graduates; postgraduate. These 5 groups have been clustered into 3 to generate 3 educational groups consistent through the three countries. The substantial results do not change: the residual increases for more-educated workers and decreases for less-educated ones.

Table 3.3 compares the educational characteristics of natives and immigrants in the labour force for men and women in the US and UK. A couple of findings immediately emerge from the table: compared to the UK, the US labour force is on average characterised by having a small share of low-educated workers. Considering natives in 2008, only 5% and 4% for men and women, respectively, were classified in the low-education group; while for the UK counterparts, these shares increase to 27% (men) and 31% (women). More differences arise in the US when comparing less educated natives with equivalent immigrants: as the existing literature documents, most immigrants to the US are relatively less educated. In 2008, 29% of male immigrants in the US had a lower level of education; this share is 13% points higher than the share of male immigrants in the UK. The situation is better for female immigrants in the US; in fact, 18% of them have a low level of education. This result is very similar for female immigrants in the UK (0.16).

In contrast, the mix of immigrants to the UK has become much more educated over time compared not only to immigrants in the US, but also to natives UK. In 2008, almost half of the immigrants, both men and women, to the UK were highly educated, compared to 25 and 32 percent for male and female natives, respectively. This gap is less evident when comparing higher educated natives and immigrants in the US, where 33% of native males have a degree or postgraduate degree compared to 29% of immigrants; there is even less of a difference between native and immigrant women.

3.5.2 Wage Dispersion for Immigrants and Natives

The pattern and the evolution of wage dispersion in the two countries over the same period are shown in Figures 3.1a to 3.2d, together with Tables 3.4a and 3.4b; the results are displayed for all workers as well as immigrants and natives separately.

Figures 3.1a and 3.1b use as a measure of the dispersion of wage inequality the standard deviation of the log hourly real wage of workers. The first remarkable fact is that for all workers, represented by the solid line in the graphs, the levels of wage inequality in the US and the UK are fairly similar, though

slightly higher for men in the US[40] For women, at the beginning of the sample year, wage dispersion was higher in the UK than it was in the US, but in 2008 the situation was reversed, with an increase for women in the US and a decrease for women in the UK.

Comparing immigrants to natives, both figures show that, on average, the former experience higher levels of wage dispersion. In particular, immigrants in the UK experience greater dispersion than their counterparts in the US. The reverse is true for women, where, except for the years 1996 and 1997, female immigrants to the UK experienced less inequality than their counterparts in the US.

These results are confirmed by Tables 3.4a and 3.4b summarising different measures of wage dispersion and their main trends over time. Table 3.4a compares the standard deviation and 90-50 and 50-10 gaps between the US and UK, considering men, women, all workers, natives and immigrants separately. The top panel reports wage dispersion for men in both countries. In 1994, wage inequality for all male workers was 0.547, and this value increased to 0.569 in 2008. Similarly, in the UK, standard deviation increased from 0.547 to 0.559; the standard deviation of wages was higher for immigrants in the UK in 1994 (0.612) than for those in the US (0.600). However, at the end of the sample period, immigrants in the UK experienced a decrease to 0.608 while, following the trend for all workers, wage dispersion for immigrants in the US increased. The trends go in opposite directions for females in the two countries; in the US, wage inequality increased for all female workers (from 0.509 to 0.525) and immigrants (0.552 to 0.580), while it decreased for both all workers (from 0.542 to 0.504) and immigrants (0.578 to 0.531) in the UK.

The same table also reports a measure of dispersion by looking at workers located in the upper and lower tails of the wage distribution. In general, wage

[40] However the similar measure of dispersion used in chapter 2 documented a slightly different picture when comparing the UK and the U.S, namely the level of wage inequality in the US was higher than that of the UK. This difference may be first of all related to the different time period analysed (between 1987 and 2004) and to the different data used, (MAY/ORG CPS for the U.S. and the GHS for the UK). Lemieux (2006) and Autor et al. (2008) pointed out how the measurement error of data could lead to different levels of wage inequality. In addition the different in sample size could also affect the precision of the estimates; for example the number of observations for men in the GHS data ranges between 1,650 and 4,400 while for the same group in the LFS data they range between 12,000 and 32,000. Finally the different levels in the wage dispersion could be also due to the different approaches used in this work and in that of Lemieux (2002) to address the top-coding issue fot the US data.

dispersion for workers located in the upper tail of the wage distribution increased over time for all natives and immigrants, both male and female, with the exception of females in the UK: in this case, dispersion of workers in the upper tail of the wage distribution decreased for both natives and immigrants, though the latter suffer from more inequality. Considering the 90-50 and 50-10 wage gaps for male workers in the US, the 90-50 log hourly wage gap increases over the sample period for all workers and for both men and women; similarly, for workers located on the lower tail of the wage distribution, wage dispersion decreased from 1994 to 2008, but immigrants experienced the largest decrease. Female workers located on the lower tail of the wage distribution did not display a decrease in wage dispersion; in fact, there was a slight increase, which was higher for female immigrants. For the UK, the 90-50 log hourly wage gap has a different sign for men versus women; the former experience an increase in wage dispersion, which is larger for immigrants; the latter are instead affected by an overall decrease in wage inequality, except for the increase for female immigrants. More similarities are displayed for workers located in the lower tail of the wage distribution: in all cases, wage dispersion decreases, and the largest decrease is found for male immigrants. A more uniform pattern is observed when considering workers located in the lower tail of the wage distribution: in all cases (men and women, immigrants and natives in the US and the UK), wage dispersion has decreased over time.

The same table also shows what happens to wage dispersion when foreign-born workers are not counted in the sample: when considering natives only, in all cases, the level of wage inequality is slightly lower, though still increasing. This means essentially that the presence of immigrants in the labour force does not strongly affect the level of wage inequality experienced by the countries, with the trends being the same.
Table 3.4b summarises changes over the sample period: immigrants located in the lower tail of the wage distribution experience the largest decrease in wage dispersion; this corresponds to a 0.1 log percentage point decrease for male immigrants in the US and 0.147 in the UK. Similar trends of smaller magnitudes occur for female immigrants.

Figure 3.2a to 3.2d plot the kernel distribution of log hourly wages for natives and immigrants in both countries for the years 1994 and 2008. The figures

clearly show the difference in the wage distributions between natives and immigrants. Particularly in the US and for male workers, the density of log hourly wage for immigrants is left-shifted compared to natives. The two plots for women in the U.S. are qualitatively similar, though the gap in the density distribution is smaller in 2008. Results for the UK show a substantial overlap in wage densities for women in 2008, while in 1994 the density for women born abroad appears to be slightly right-shifted compared to natives. This pattern is similar for male workers in the UK in 1994, though the mean log hourly wage for natives is higher than for immigrants; in 2008, consistent with the US, the wage density of immigrants in the UK is left-shifted compared to their native counterparts.

Figures 3.4a and 3.4b plot levels and trends using as a measure of dispersion the variance of the residual[41] hourly wage for men and women in the US and the UK. The residual variance has been calculated separately from three different regressions, respectively, for the whole labour force, natives and immigrants. Confirming previous results, these graphs show that the dispersion in the wages of immigrants is higher[42] than it is for natives, especially for the UK, with the bulk of the increase for foreign-born immigrants working in the UK occurring at the end of the 2000s when immigration was rising strongly.

This is particularly true for males, while there is more of a convergence in level and pattern for males in the US: starting in 2002, there is not much difference in residual dispersion between immigrants and natives. More variation remains for immigrant females in the US, and even more for the UK.

When treating immigrants and natives as two separate groups, interesting differences arise: in all cases, the level of wage dispersion in the residual for natives decreases, but the almost unchanged trends suggest that had immigration not occurred, the residual wage inequality of natives in both countries would have increased anyway, albeit at a slightly lower level. This description implies that

[41]Wage residuals are obtained from a series of regression models fit separately by gender, immigrant status and year. The models include controls for education, experience, and interaction terms. When including a control for country of origin, the main results do not change too much. On average, the residual obtained when adding the control for area of origin is 0.002 lower than the one obtained when controlling for human capital variables only.

[42] One concern related to the increase in the share of immigrants is the area of origin; particularly, the UK dispersion in residual could also be due to different inflows of immigrants. To address this issue, I also control for area of origin both in the original regression and in the logit model; the residual variance decreases a bit, mainly for immigrants, while results for all workers and natives do not change significantly. Nevertheless, the distance between residual variance of immigrants and natives still remains.

despite the changing in composition that have characterised immigrants in the past decades, the presence of immigrants cannot be considered to be even partially responsible for the increasing level of residual wage inequality in both the US and the UK between 1994 and 2008. In a similar vein to the recent interpretation of Card (2009), this suggests that the role of immigration on the increase of inequality in the workforce is almost negligible showing no effect on the inequality of native wages. Figures 3.3a to 3.3d plot the variance and the 90-50 residual of log wages for all workers (men and women, respectively) together with the share of immigrants in the labour force. As depicted in Figure 3.3a, the level of residual wage inequality has been increasing a little over the sample years, with a somewhat larger increase during the most recent years of the sample. On the other hand, both in the US and UK, the share of immigrants has been increasing continuously over this period. The rate of increase in immigration is larger in the UK, although the share of immigrants remains higher in the US. The results are qualitatively similar for women, except that female immigrants in the UK are more numerous than males and even more numerous than those in the US. Figures 3.3c and 3.3d compare the trends between the 90-50 residual gap and the share of immigrants over the sample years. As has been shown in previous studies and in Chapter 2, an increase in wage inequality occurred particularly on the upper tail of the wage distribution.

3.5.3 Residual Inequality, the Composition Effect and the Increase in Immigration

Table 3.5a presents the decomposition of total wage variation into different components (between- and within-variance), analysing all workers, natives and immigrants separately.

Between 1994 and 2008, the change in wage variance was greater in the US than it was in the UK, particularly for men. Over the same time period women in the UK experienced a decrease in the total wage variance. In all cases over the period analysed, the residual component accounted for most of the variation observed in the total wage variance.When looking at natives only, the total residual wage change decreases, though for a small amount, with respect to the total residual wage change observed when looking at all workers. However the actual residual variance for native women in the UK decreases by about one third

with respect to the actual residual variance of women in the whole labour force (natives and immigrants).

In a similar vein, Table 3.5b presents the results for natives and all workers when holding fixed the characteristics of workers as well as the share of immigrants in the labour force. On average, when the characteristics of the labour force are held fixed at their 1994 levels, the level of residual wage inequality decreases on average for both men and women in the US and UK, confirming the main results of Lemieux. However, when the share of immigrants remains fixed at the level of the base year (1994) in addition to fixing the characteristics of the labour force, an additional decrease in the residual wage inequality occurs suggesting a very small effect of the increasing share on immigrants on the residual wage inequality.

Figures 3.5a to 3.10d present the results of the counterfactual reweighting approach proposed by Lemieux that helps to account for the roles of composition effects and the share of immigrants in changes in residual wage inequality. The figures compare, for each country, for men and women separately, for all workers and for natives only, the actual residual variance from 1994 to 2008 to the counterfactual variance that would have existed if the distribution of skills had remained at the same level of the initial (1994) or the final (2008) year. The composition effect is represented in the figures by the distance between the actual variance and the counterfactual variance (residual holding skills constant at its 1994 or 2008 level).

As displayed in Figures 3.5a and 3.5b, when the distribution of skills characteristics of the labour force is held fixed at its 1994 level, the remaining growth of the residual variance accounts for a very small share. The results hold both for men and women in the US and the UK. This evidence confirms the findings of Lemieux (2006) but now analysing natives and the whole labour force separately. However, the role of the composition component becomes less significant once the share of immigrants is also held fixed at its 1994 level.

On average, between 1994 and 2008 when the level of education and experience of the labour force and the share of immigrants is held fixed at a base year, the change in residual variance only accounts for about 1 to 2% of the change in total residual variance for the same time period; consistent with Card (2009) the effect is very small but not causal.

Figures 3.7a to 3.8b show the results when the reweighted approach is restricted to natives only; this should yield the counterfactual residual variance for natives if the share of immigrants and the level of education and experience of natives remained constant in the base year; the figures prove that the pattern of residual wage inequality observed for natives in both the US and UK is mainly due to changing characteristics of natives, while the increased supply of immigrants does not account for much of the inequality.

3.5.4 Residual Wage Inequality: Upper and Lower Wage Gaps

During the 1990s, wage inequality occurred mainly in the upper tail of the wage distribution. This section examines alternative measures of wage inequality, specifically the 90-50 and 50-10 residual gaps, comparing the actual residual to the computed residual holding distribution of skills and share of immigrants at their 1994 levels to understand how much change in the residual and composition affect all workers, natives only and immigrants only at each tail of the distribution.

Looking at the evolution of the 90-50 gap, Figures 3.9a to 3.9d show that residual wage dispersion for workers located on the upper tail of the wage distribution increased for the workforce in both countries for both men and women. When, in addition to holding fixed the characteristics of workers in the labour force, the share of immigrants is held constant at its 1994 level, the role of the residual component in the upper tail of the wage distribution decreases even more. However, there is little effect of the share of immigrants when considering natives only. Results are qualitatively similar for women.

Looking at the lower tail of the wage distribution, the evolution of the residual variance is similar across groups (all workers, natives) and between the two countries; in fact, the 50-10 residual gap is characterised by a general decrease over time. Consistent with the results of the previous section, the effects of composition appear to be more important: changes in the share of foreign-born workers exert almost no effect. This suggests that the presence of immigrants in the labour force not only has little effect on the change in and level of residual

wage inequality, but it also does not have a large effect on the distribution of wages in both the US and UK.

3.6 Conclusion

This paper addressed empirically the question of whether increasing immigration in the US and the UK could have been associated with the rising residual wage inequality observed between 1994 and 2008. As far as I am aware, except for a recent study by Card (2009), almost no attention has been devoted to the role of immigrants' unobservable skills in the context of wage inequality. By applying Lemieux's methodology (2006), this paper controls for changing characteristics of workers and adds the immigration dimension to control for the increasing presence of foreign-born workers in the labour force. In line with Card's recent findings, the results show that the presence of immigrants does not have a causal relation with the increase in residual inequality. Even when removing immigrants from the labour force, the trend of residual variance does not differ much from the trend observed when including both natives and immigrants in the samples. This suggests that the increasing share of immigrants in both the UK and the U.S. labour force occurred between 1994 and 2008 cannot help (much) to account for the increasing residual wage observed in both countries over the same time period. Therefore changes in composition of the labour force account for most of the residual wage inequality.

However, when the share of immigrants in the labour force is held fixed at its value in the base year, the change of the residual variance between 1994 and 2008 ranges from -0.001 to 0.008 log points for all workers (men and women) in the U.S. compared to the change of the total residual variance for the same group varying between 0.016 and 0.024 log points; the similar figure for all workers (men and women) in the UK shows that the change in the residual variance obtained by holding fixed at a base year both characteristics all workers and the share of immigrants in the labour force, ranges between -0.006 to 0.009 log points compared to the change of the total residual variance ranging from -0.015 to 0.016. Although the composition effect still explains most of the growth in the residual for natives and the whole labour force, the results also suggest that the presence of immigrants plays a role in explaining the growth in the residual variance observed, albeit a very small share. On average, the non-causal effect of

immigration on residual variance ranges between 1% and 5% of the observed change between 1994 and 2008.

This finding suggests that the presence of immigrants in the labour force not only has little effect on the change and level of residual wage inequality, but it also does not affect the distribution of wages in either the US or the UK. In terms of public policy, the results imply that the inflow of workers from abroad has no negative effect on natives' wages in the receiving countries, confirming that the effect of immigration on wages should not be a concern in the public policy agenda.

Table 3.1: Proportion of Immigrants in the Labour Market

	Men	Women
US		
1994	0.10	0.08
2008	0.15	0.12
UK		
1994	0.06	0.07
2008	0.12	0.12

Notes: The table shows the proportion of immigrants in the US and UK labour force. Samples include people aged 16 to 59 for women and 16 to 64 for men with positive potential experience; working full time, full year, employed only s in their main job only.

Table 3.2: Areas of Origin of Immigrants to the United States and Great Britain, 1994-2008

	1994	2008
U.S.		
Europe	0.13	0.11
Asia	0.25	0.25
Africa	0.14	0.031
Americas	0.46	0.58
Oceania	0.003	0.005
Other	0.014	0.028
UK		
Europe	0.39	0.51
Asia	0.26	0.20
Africa	0.16	0.17
Americas	0.13	0.07
Oceania	0.05	0.03
Other	0.02	0.02

The table shows the areas of origin of immigrants in the US and UK labour force. Samples include individuals in the labour force with positive potential experience; working full time, full year, as employees in their main

Table 3.3: **Educational Attainment of Natives and Immigrants, US and UK**

Level of Education	**Natives**		**Immigrants**	
	Men	Women	Men	Women
A. **US**				
Lower				
1994	0.08	0.06	0.30	0.22
2008	0.05	0.04	0.29	0.18
Intermediate				
1994	0.64	0.67	0.43	0.51
2008	0.62	0.60	0.42	0.47
Higher				
1994	0.28	0.27	0.27	0.27
2008	0.33	0.36	0.29	0.35
B. UK				
Lower				
1994	0.28	0.43	0.21	0.24
2008	0.27	0.31	0.16	0.16
Intermediate				
1994	0.55	0.30	0.37	0.34
2008	0.48	0.37	0.24	0.25
Higher				
1994	0.17	0.17	0.42	0.41
2008	0.25	0.32	0.56	0.59

The table shows the educational characteristics of natives and immigrants in the US and UK labour force. The lower education group includes workers who have completed compulsory education, i.e., less than a lower secondary education (from 0 to 11 years of schooling). The intermediate category includes workers with qualifications that exceed those of a high-school dropout but do not reach those of a college-degree holder (both excluded), this corresponds to any individual with years of schooling at least 12 and at most 15 years of schooling. The higher education group refers either to individuals with graduate or postgraduate education, corresponding to 16 or more years of schooling.

Table 3.4a: Wage Dispersion of Log Hourly Wage, US and UK

	US			UK		
	Standard deviation	90-50	50-10	Standard deviation	90-50	50-10
Males						
A. All workers						
1994	0.547	0.684	0.721	0.545	0.706	0.630
2000	0.543	0.724	0.674	0.549	0.738	0.616
2008	0.569	0.765	0.687	0.559	0.751	0.620
B. Natives						
1994	0.537	0.674	0.697	0.541	0.703	0.627
2000	0.529	0.693	0.694	0.544	0.734	0.610
2008	0.556	0.754	0.693	0.552	0.730	0.617
C. Immigrants						
1994	0.600	0.876	0.693	0.612	0.770	0.783
2000	0.581	0.917	0.573	0.602	0.830	0.677
2008	0.613	0.976	0.593	0.608	0.861	0.636
Females						
A. All workers						
1994	0.509	0.693	0.598	0.542	0.689	0.598
2000	0.498	0.688	0.553	0.502	0.668	0.584
2008	0.525	0.722	0.597	0.504	0.671	0.558
B. Natives						
1994	0.503	0.683	0.591	0.522	0.681	0.601
2000	0.490	0.664	0.580	0.500	0.666	0.573
2008	0.513	0.701	0.607	0.500	0.667	0.554
C. Immigrants						
1994	0.552	0.787	0.580	0.578	0.690	0.661
2000	0.543	0.870	0.518	0.514	0.650	0.659
2008	0.580	0.896	0.568	0.531	0.714	0.592

Notes: The table reports measure of dispersion of the log hourly wage for natives, immigrants male and females in the US and the UK. Data are drawn from ORG/CPS and LFS for the US and UK respectively.

Table 3.4b: Changes in Wage Inequality in the US and UK, 1994-2008

	UK			US		
	Standard Deviation	90-50	50-10	Standard Deviation	90-50	50-10
Males						
A. All workers						
1994-2008	0.022	0.081	-0.034	0.014	0.043	-0.018
B. Natives						
1994-2008	0.019	0.08	-0.004	0.011	0.026	-0.010
C. Immigrants						
1994-2008	0.013	0.1	-0.1	0.013	0.089	-0.147
Females						
A. All workers						
1994-2008	0.016	0.029	-0.001	-0.015	-0.018	-0.037
B. Natives						
1994-2008	0.008	0.037	-0.011	-0.017	-0.014	-0.044
C. Immigrants						
1994-2008	0.028	0.108	-0.011	0.009	0.027	-0.054

Notes: The table shows changes in wage dispersion for the log hourly wage of male and natives, natives and immigrants, in the US and the UK between 1994 and 2008

Table 3.5a: Wage Decomposition: Natives, Immigrants and all Workers

	1994-2008 UK			1994-2008 US		
	Natives	Immigrants	All workers	Natives	Immigrants	All workers
A. Men						
Actual residual	0.021	0.025	0.028	0.018	0.005	0.017
Predicted value	-0.012	-0.023	-0.012	0.002	0.011	0.007
Total wage variance	0.012	0.016	0.016	0.020	0.016	0.024
B.Women						
Actual residual	0.003	0.033	0.010	0.008	0.006	0.009
Predicted value	-0.025	-0.031	-0.025	0.004	0.026	0.007
Total wage variance	-0.018	-0.009	-0.015	0.012	0.032	0.016

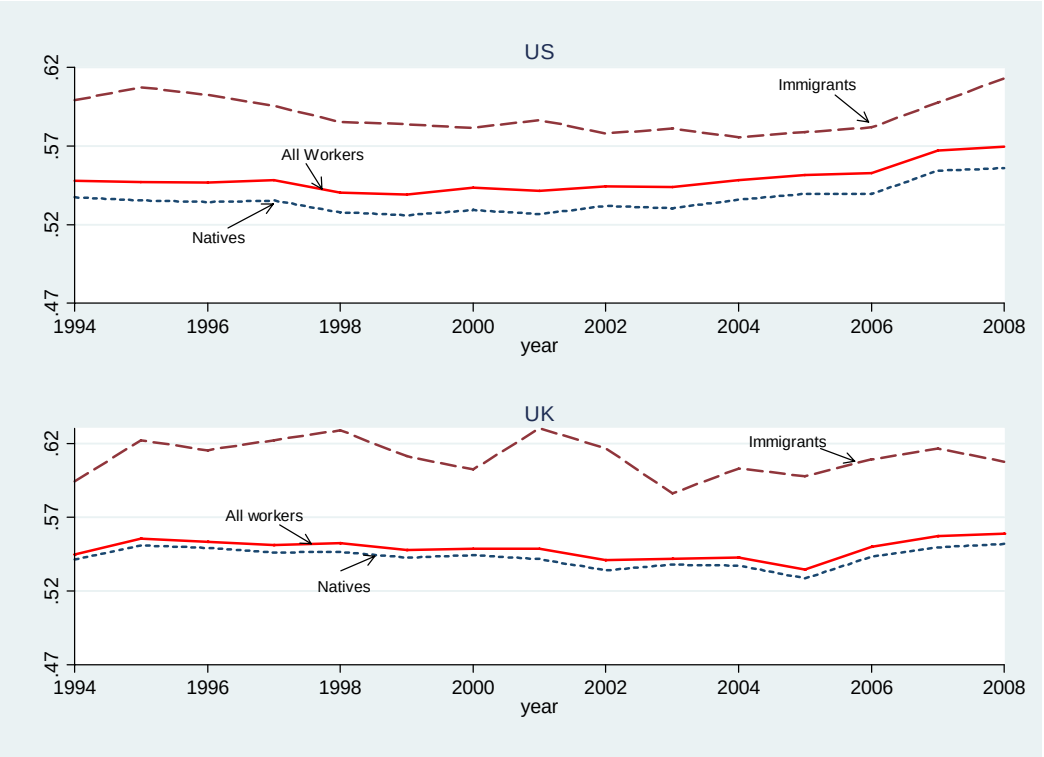
Notes: Residual wage variance is based on standard Mincer wage equation, fit separately by year, gender, and immigration status.

Table 3.5b: Wage Decomposition for Natives and all Workers
(Holding Skills and Share of Immigrants at their 1994 Level)

	UK 1994-2008		US 1994-2008	
	Natives	All workers	Natives	All workers
A. Men				
Actual Residual	0.021	0.028	0.018	0.017
Skills in 1994	0.006	0.006	0.010	0.009
Skills and Share of Immigrants in 1994	0.006	0.009	0.009	0.008
Predicted Value	-0.012	-0.012	0.002	0.007
Total wage variance	0.012	0.016	0.020	0.024
B. Women				
Actual Residual	0.003	0.010	0.008	0.009
Skills in 1994	-0.009	-0.005	-0.002	-0.0003
Skills and share of Immigrants in 1994	-0.009	-0.006	0.0002	-0.001
Predicted value	-0.025	-0.025	0.004	0.007
Total wage variance	-0.018	-0.015	0.012	0.016

Notes: Residual wage variance is based on standard Mincer wage equation, fit separately by year, gender, and immigration status.

Figure 3.1a: Standard Deviation of Log Hourly Wage for Men, 1994-2008

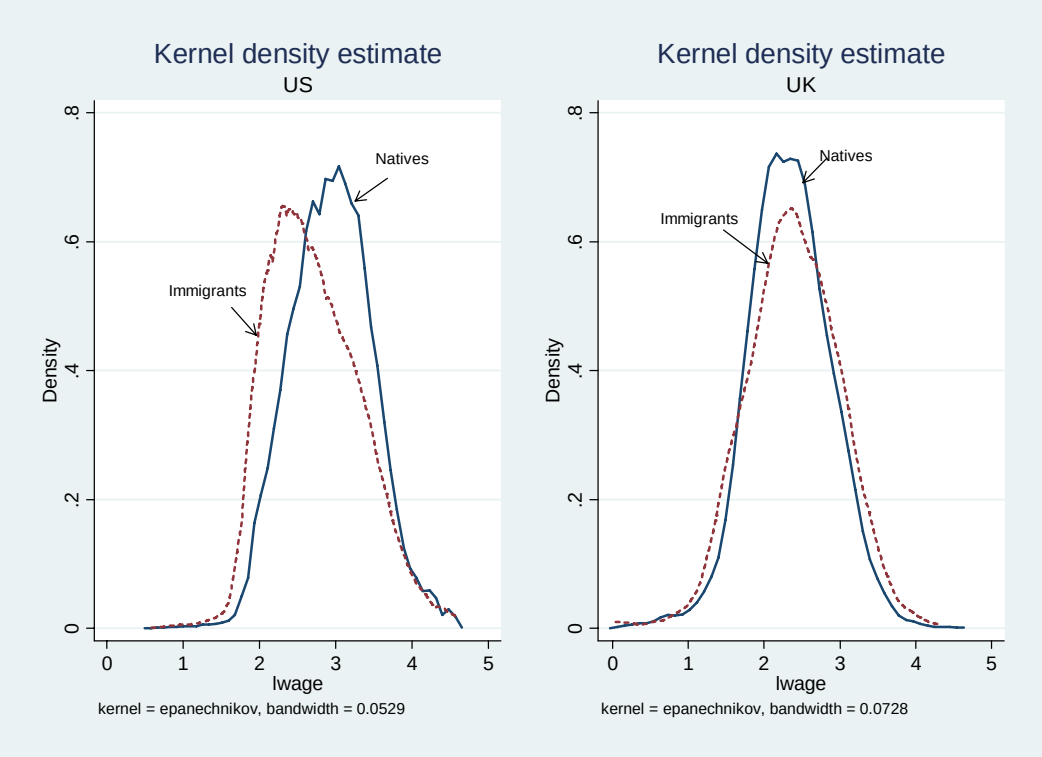


Notes: The figure plots the standard deviation of log hourly wage for male workers, separately for all workers, natives and immigrants in the US and UK from 1994 to 2008. Samples include workers in labour force age, working full time, employed only, main job with positive potential experience.

Figure 3.1b: Standard Deviation of Log Hourly Wage for Women, 1994-2008

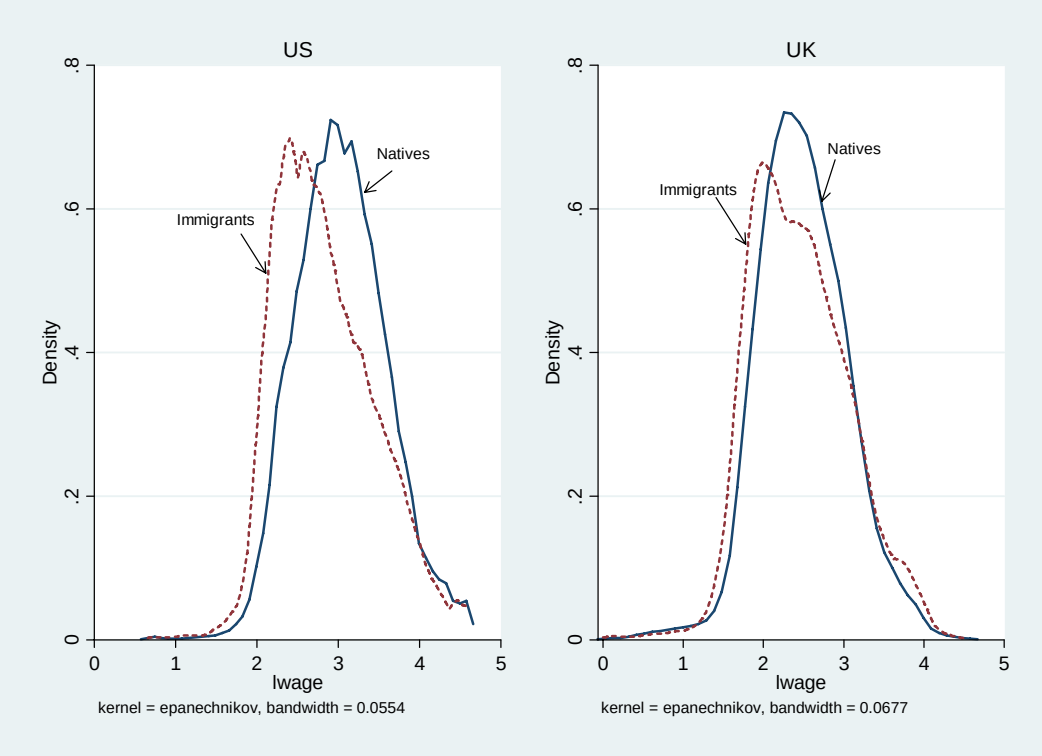
Notes: The figure plots the standard deviation of log hourly wage for female workers, separately for all workers, natives and immigrants in the US and UK from 1994 to 2008. Samples include workers in labour force age, working full time, employed only, main job with positive potential experience.

Figure 3.2a: Log Hourly Wage Distribution, Natives and Immigrants Men,1994



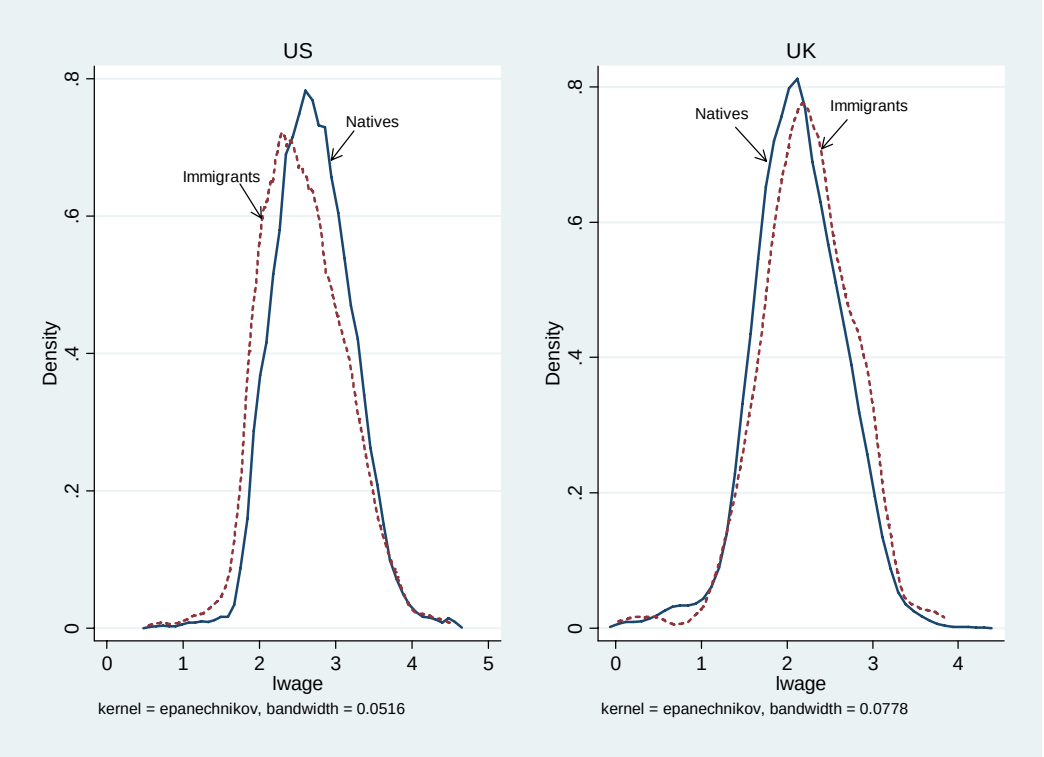
Notes: The figure plots the kernel density of the log hourly wage at 1994 for male natives and immigrants in the US and UK.

Figure 3.2b: Log Hourly Wage Distribution, Natives and Immigrants, Men 2008



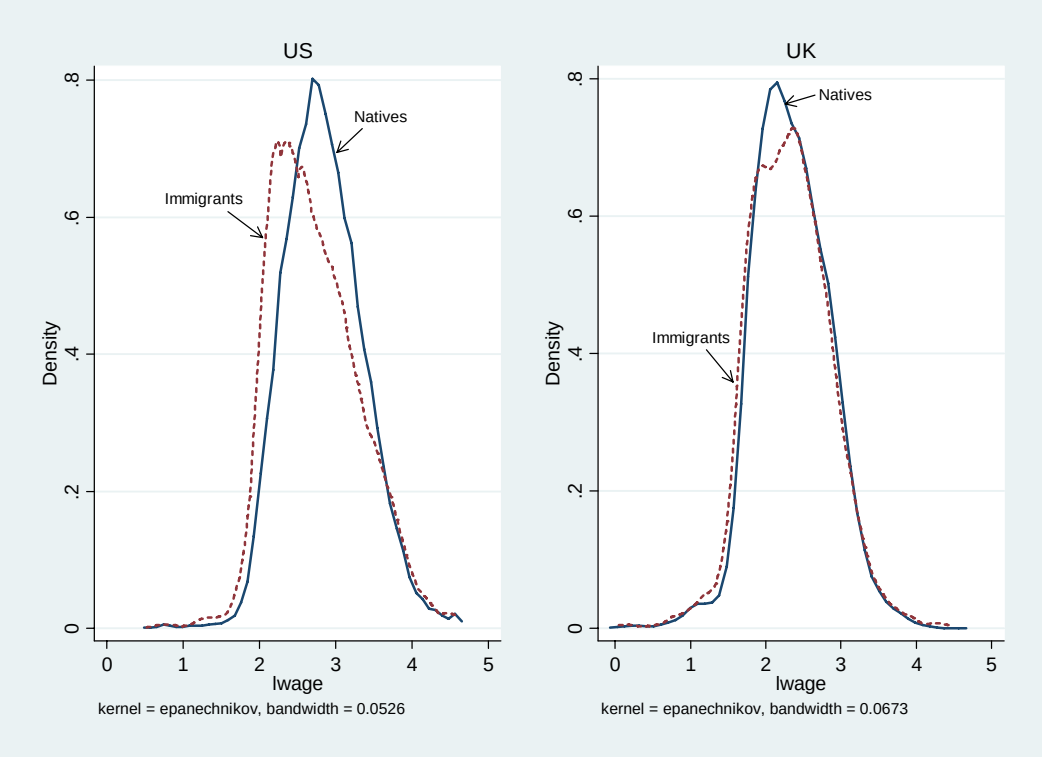
Notes: The figure plots the kernel density of the log hourly wage at 2008 for male natives and immigrants in the US and UK.

Figure 3.2c: Log Hourly Wage Distribution, Natives and Immigrants, Women1994



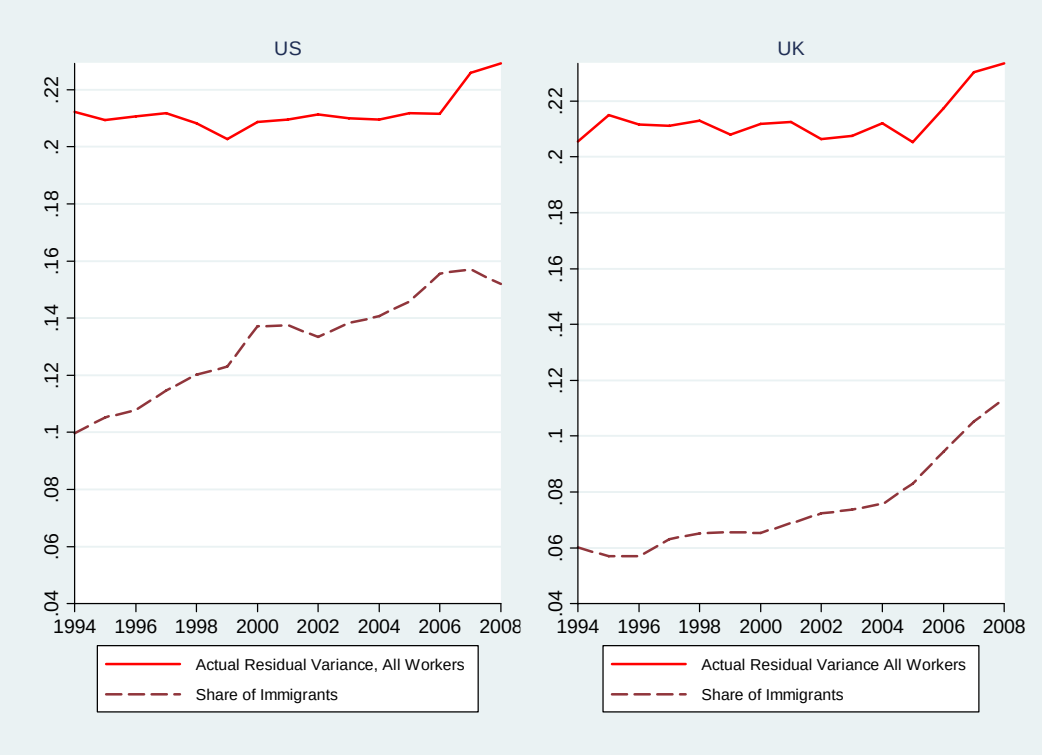
Note: The figure plots the kernel density of the log hourly wage in 1994 for female natives and immigrants in the US and UK.

Figure 3.2d: Log Hourly Wage Distribution, Natives and Immigrants, Women 2008



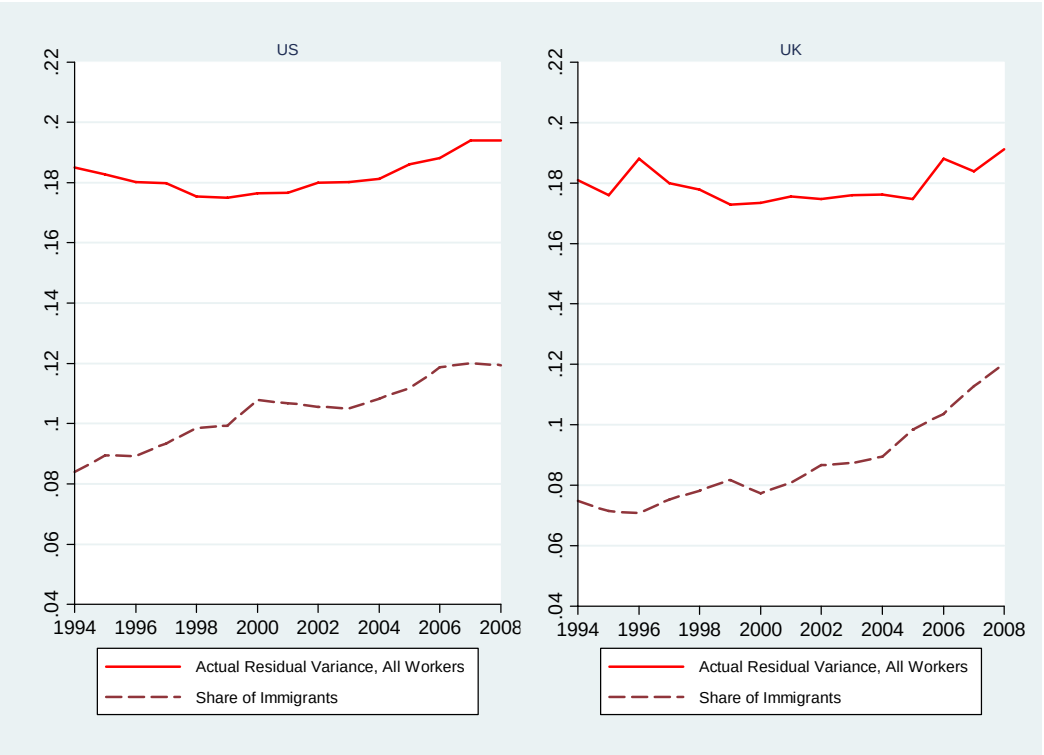
Notes: The figure plots the kernel density of the log hourly wage in 2008 for female natives and immigrants in the US and UK

Figure 3.3a: Residual Variance and Share of Immigrants, Men 1994-2008



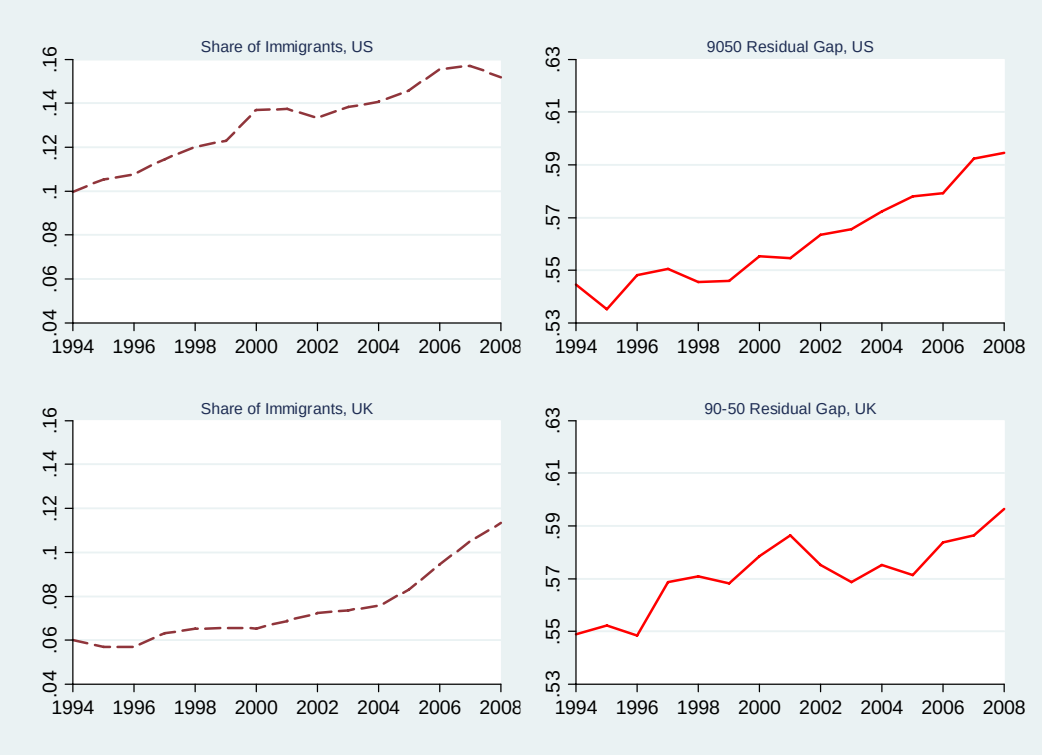
Notes: Based on ORG/CPS and LFS. The Figure plots the actual residual wage and the share of immigrants in the labour force. Hourly wages are reported in 2008 dollars and pounds. Residual wage variance is based on standard Mincer wage equation, fit separately by year, gender, immigration status.

Figure 3.3b: Residual Variance and Share of Immigrants, Women 1994-2008



Notes: Based on ORG/CPS and LFS. The Figure plots the actual residual wage and the share of immigrants in the labour force. Hourly wages are reported in 2008 dollars and pounds. Residual wage variance is based on standard Mincer wage equation, fit separately by year, gender, immigration status.

Figure 3.3c: Share of Immigrants and 90-50 Residual Gap, Men 1994-2008

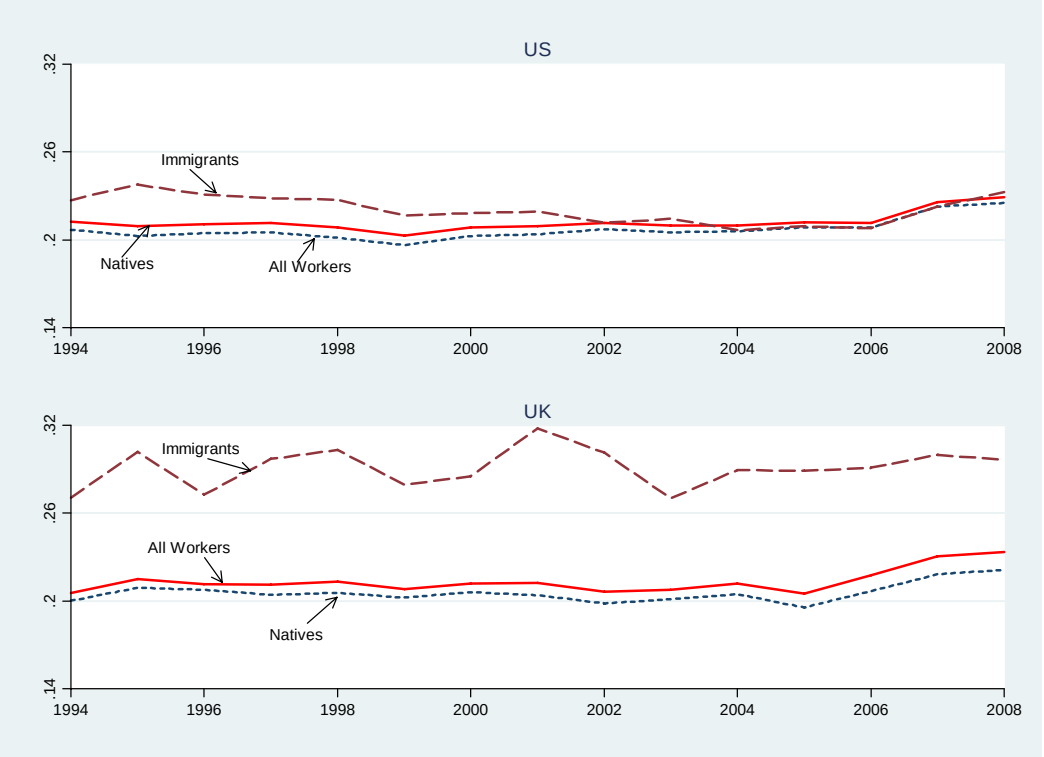


Notes: Based on ORG/CPS and LFS. Samples include men aged 16 to 64 for with positive potential experience, working full time for the full year in their main job only. The figure plots the share of female immigrants in the labour force and the 90-50 residual gap for male all workers.

Figure 3.3d:Share of Immigrants and 90-50 Residual Gap, Women 1994-2008

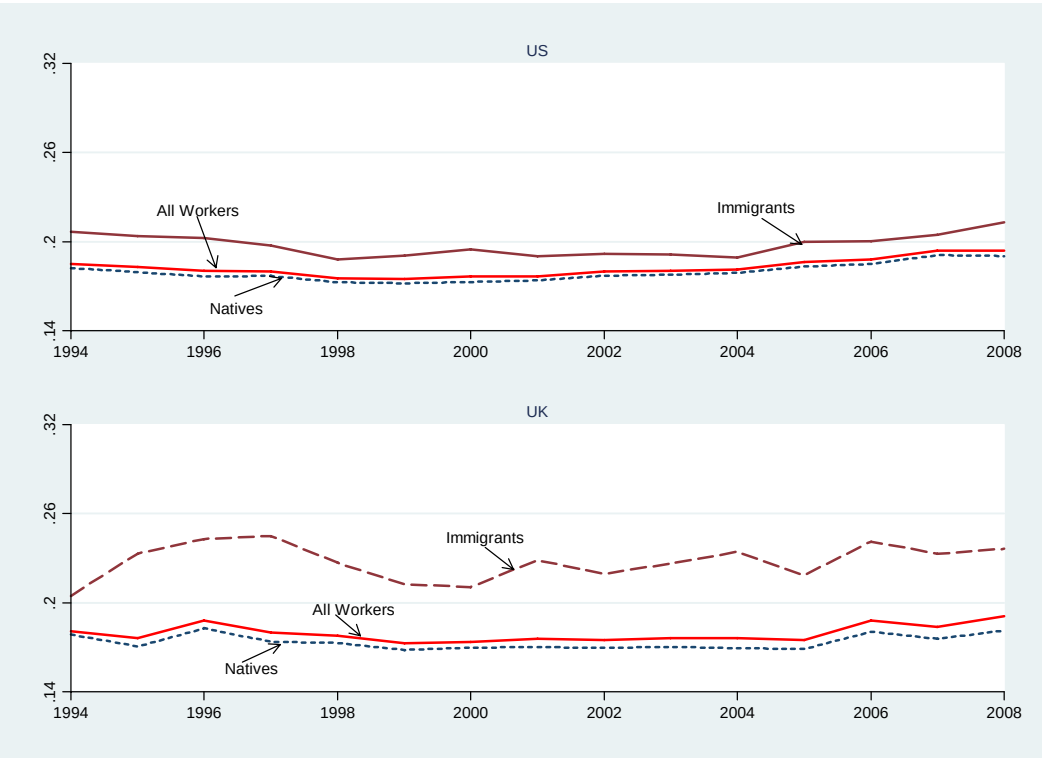
Notes: Based on ORG/CPS and LFS. Samples include men aged 16 to 64 for with positive potential experience, working full time for the full year in their main job only. The figure plots the share of female immigrants in the labour force and the 90-50 residual gap for female all workers.

Figure 3.4a: Actual Residual Variance for all Workers, Natives and Immigrants, Men



Notes: Based on ORG/CPS and LFS. Samples include workers in labour force age with positive potential experience; working full time, full year and main job only. Hourly wages are reported in 2008 dollars and pounds. Residual wage variance is based on standard Mincer wage equation, fit separately by year, gender, immigration status.

Figure 3.4b: Actual Residual variance for all Workers, Natives and Immigrants Women



Notes: Based on ORG/CPS and LFS. Samples include workers in labour force age with positive potential experience; working full time, full year and main job only. Hourly wages are reported in 2008 dollars and pounds. Residual wage variance is based on standard Mincer wage equation, fit separately by year, gender, immigration status.

Figure 3.5a: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 1994 level, Men

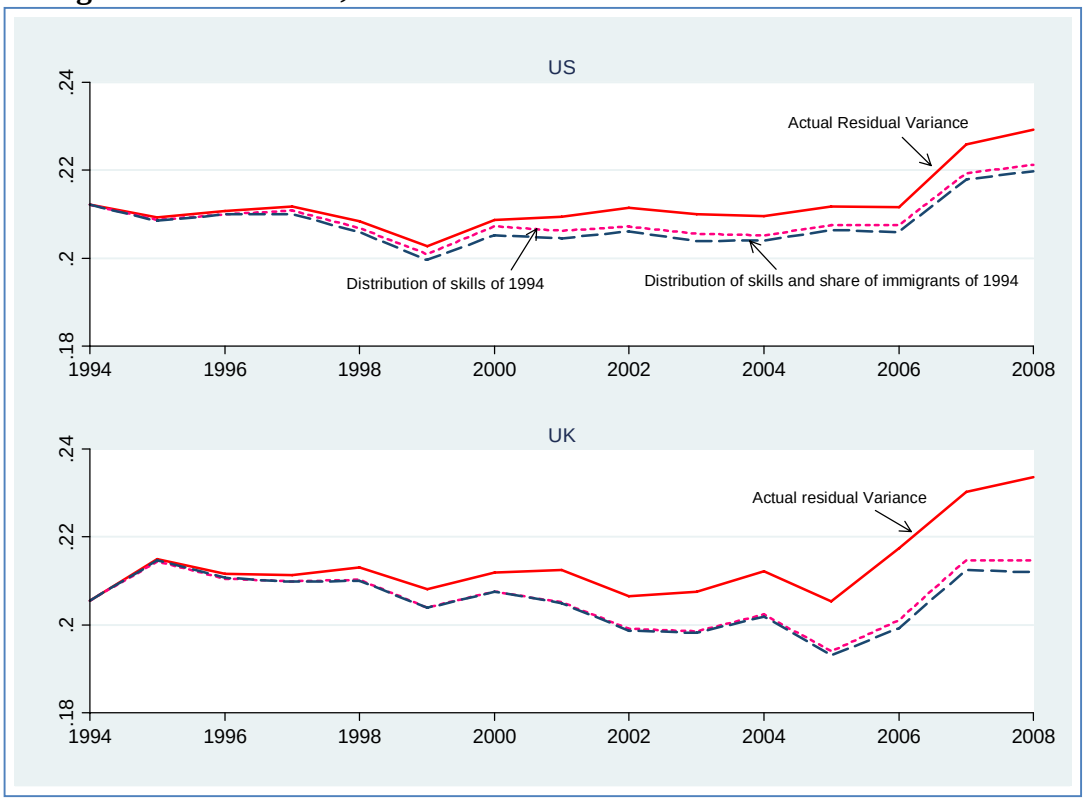
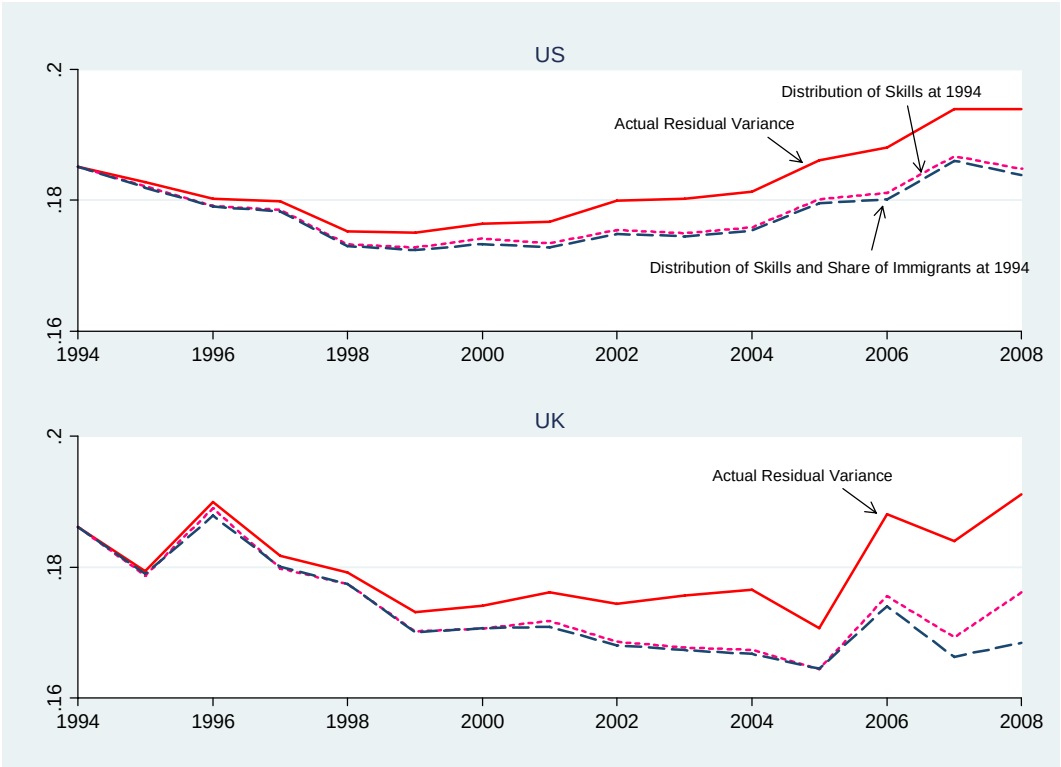


Figure 3.5b: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 1994 level, Women

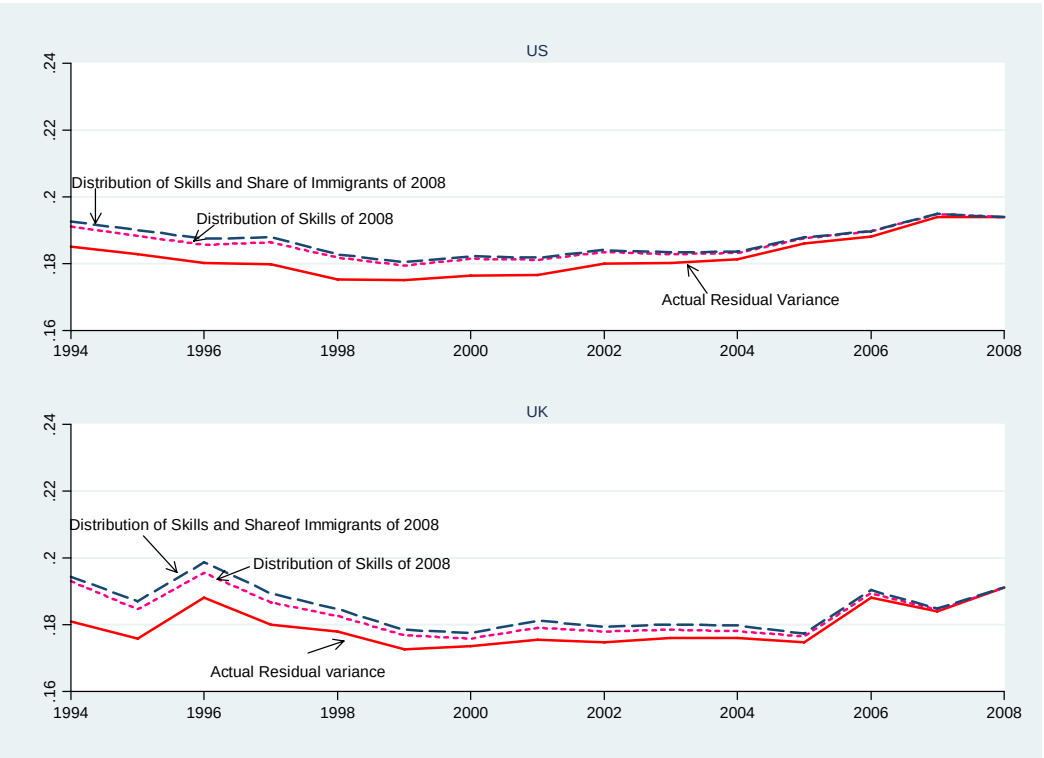


Notes: Based on ORG/CPS and LFS. Samples include all men in the labour force age with positive potential experience; working full time, full year and main job only. Hourly wages are reported in 2008 dollars and pounds. The figure plots the actual residual variance based on standard Mincer wage equation, fit separately by year,the residual holding the skills distribution in 1994 and that holding skills and the share of immigrants in 1994

Figure 3.6a: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 2008 Level, Men, Men

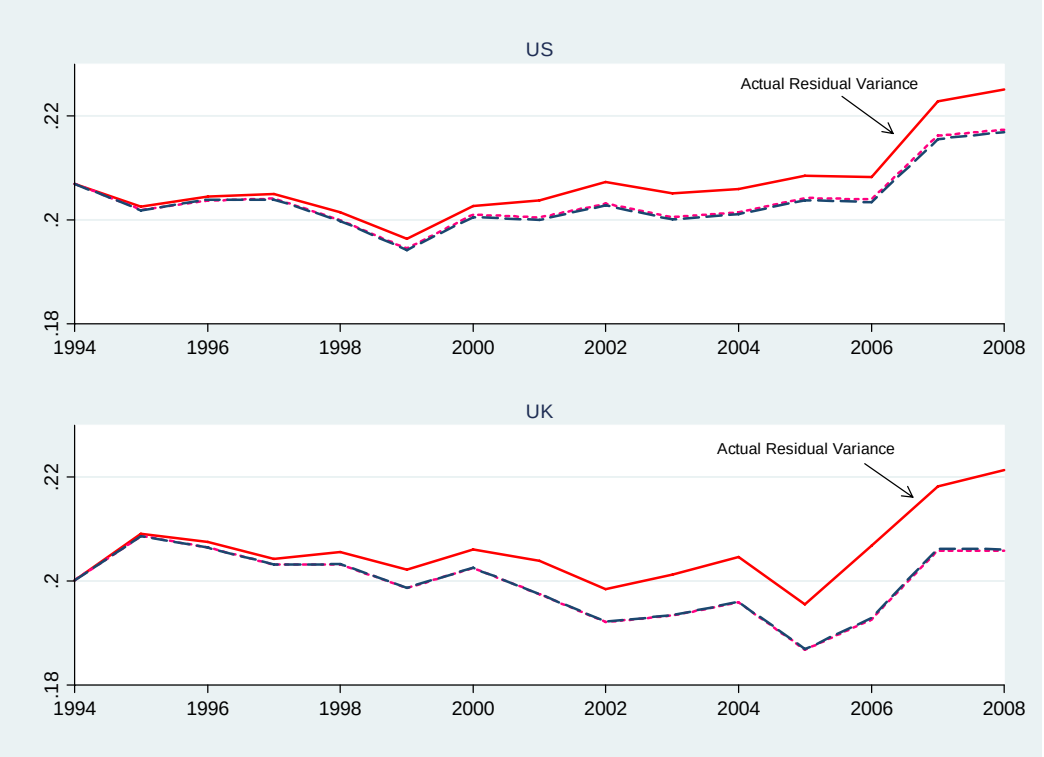
Notes: Based on ORG/CPS and LFS. Samples include men of labour force age with positive potential experience working full time for the full year in their main job only.

Figure 3.6b: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 2008 Level, Women



Notes: Based on ORG/CPS and LFS. Samples include men of labour force age with positive potential experience working full time for the full year in their main job only.

Figure 3.7a: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 1994 level, Natives Men



Notes: Based on ORG/CPS and LFS. Samples include all native men in the labour force age with positive potential experience; working full time, full year and main job only. Hourly wages are reported in 2008 dollars and pounds. The figure plots the actual residual variance based on standard Mincer wage equation, fit separately by year, the residual holding the skills distribution at 1994 and that holding skills and the share of immigrants at 1994.

Figure 3.7b: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 2008 Level, Natives Men

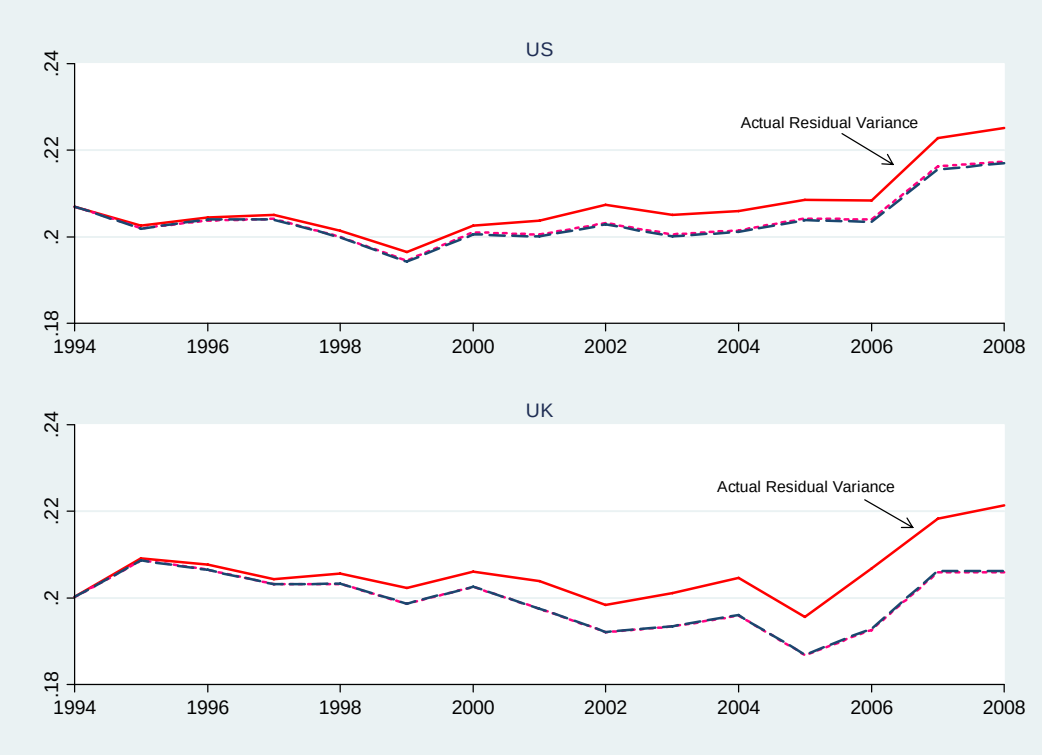
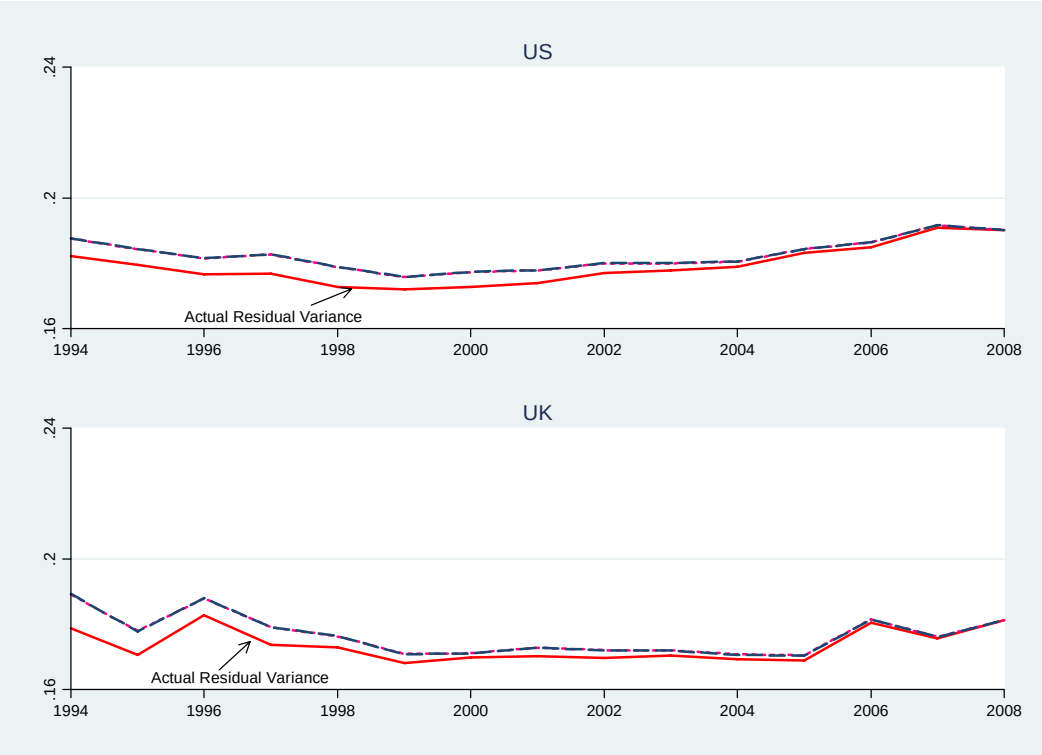


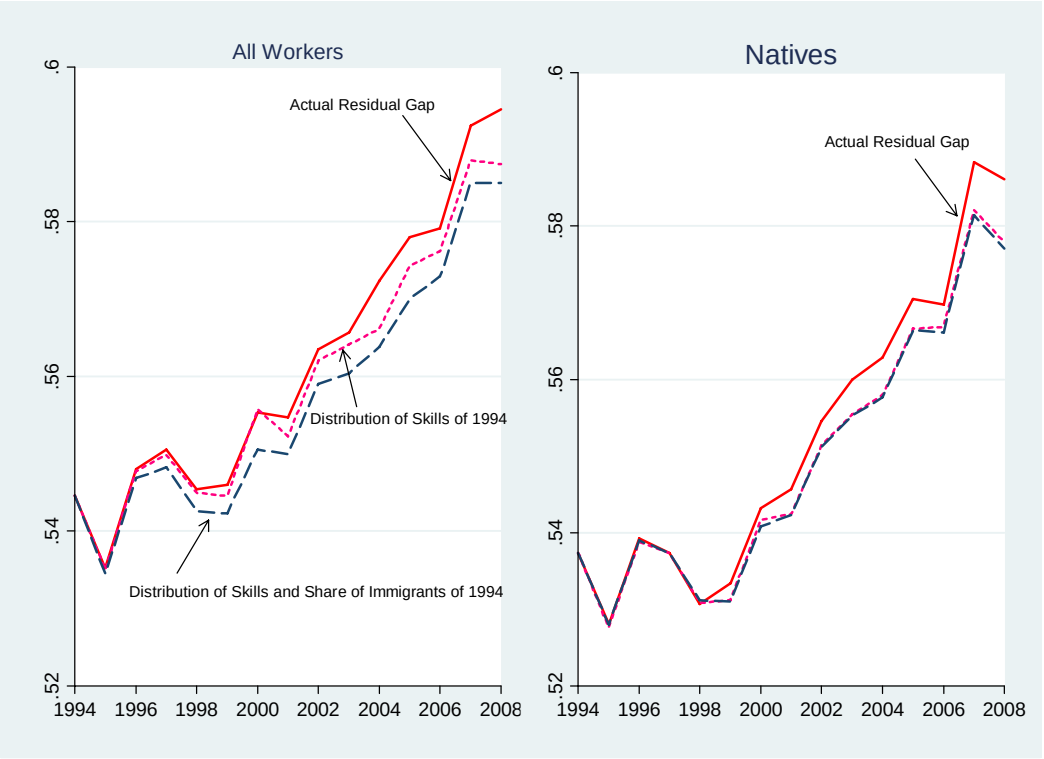
Figure 3.8a: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 2008 Level, Natives Women

Figure 3.8b: Actual Residual Variance, Distribution of Skills and Share of Immigrants at 2008 Level, Natives Women



Notes: Based on ORG/CPS and LFS. Samples include all native men in the labour force age with positive potential experience; working full time, full year and main job only. Hourly wages are reported in 2008 dollars and pounds. The figure plots the actual residual variance based on standard Mincer wage equation, fit separately by year, the residual holding the skills distribution at 1994 and that holding skills and the share of immigrants at 1994.

Figure 3.9a: 90-50 Residual Gap, Men US
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)

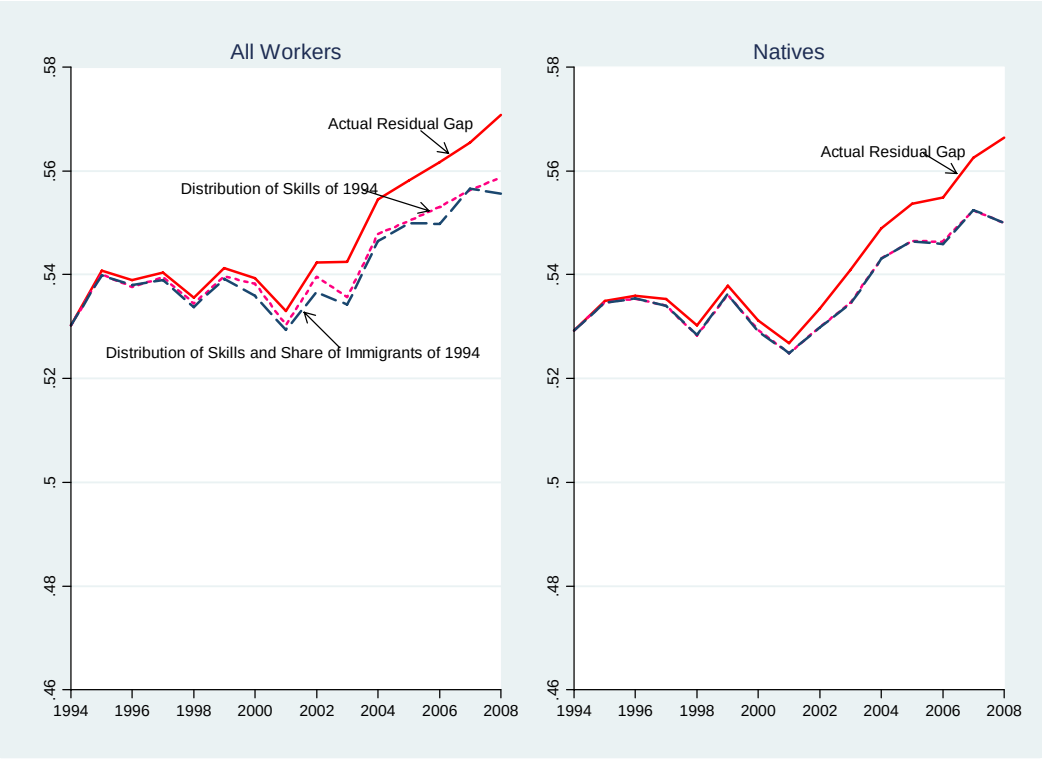


Note: The 90-50 residual gap is based on a standard Mincer wage equation fit separately by year, gender, and immigration status.

Figure 3.9b: 90-50 Residual gap, Men UK
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)

Figure 3.9c: 90-50 Residual gap, Women US

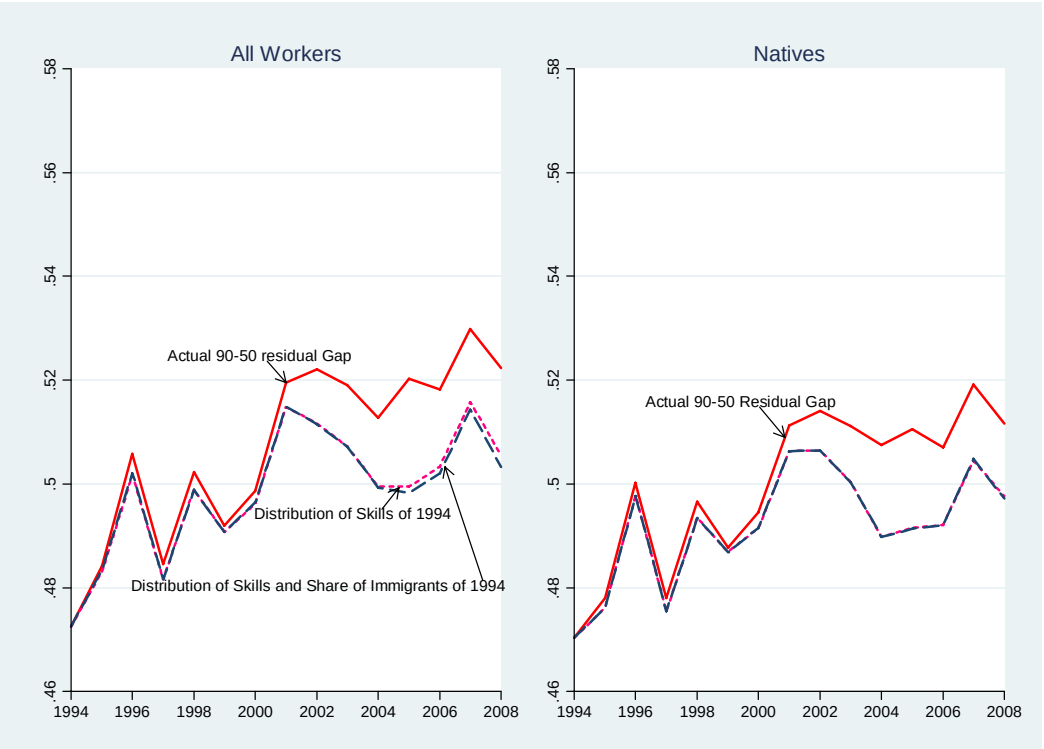
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)



Note:. The 90-50 residual gap is based on a standard Mincer wage equation fit separately by year, gender, and immigration status.

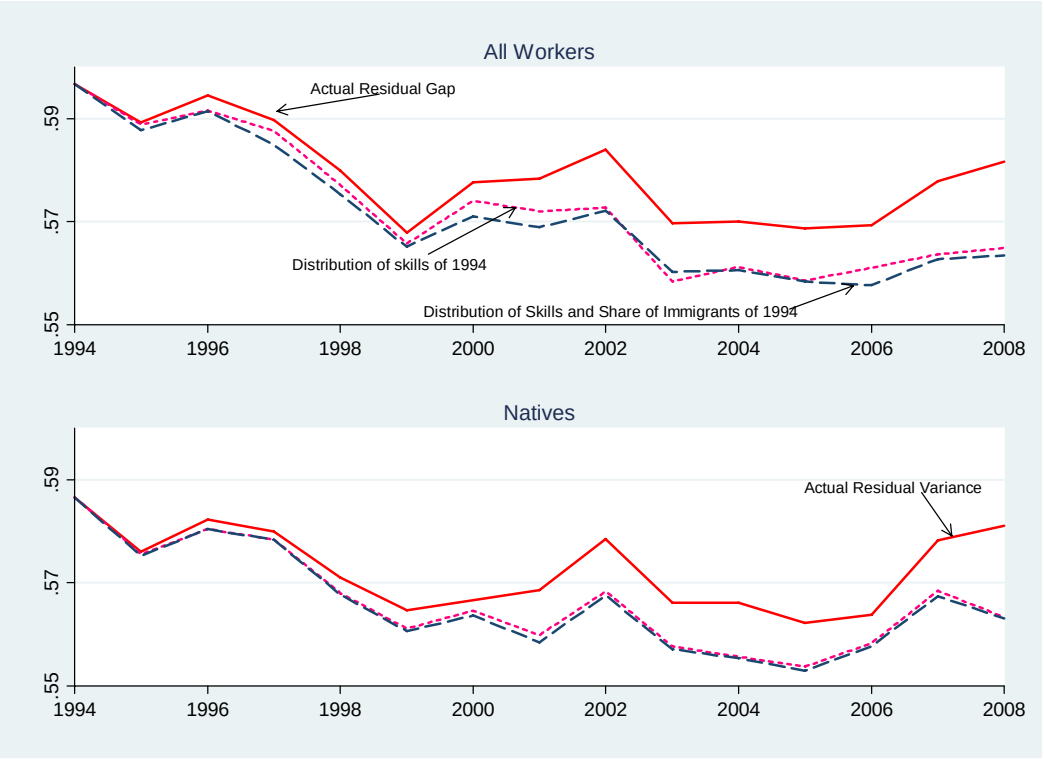
Figure 3.9d: 90-50 Residual gap, Women UK

(Holding Distribution of Skills and Share of Immigrants at 1994 Level)



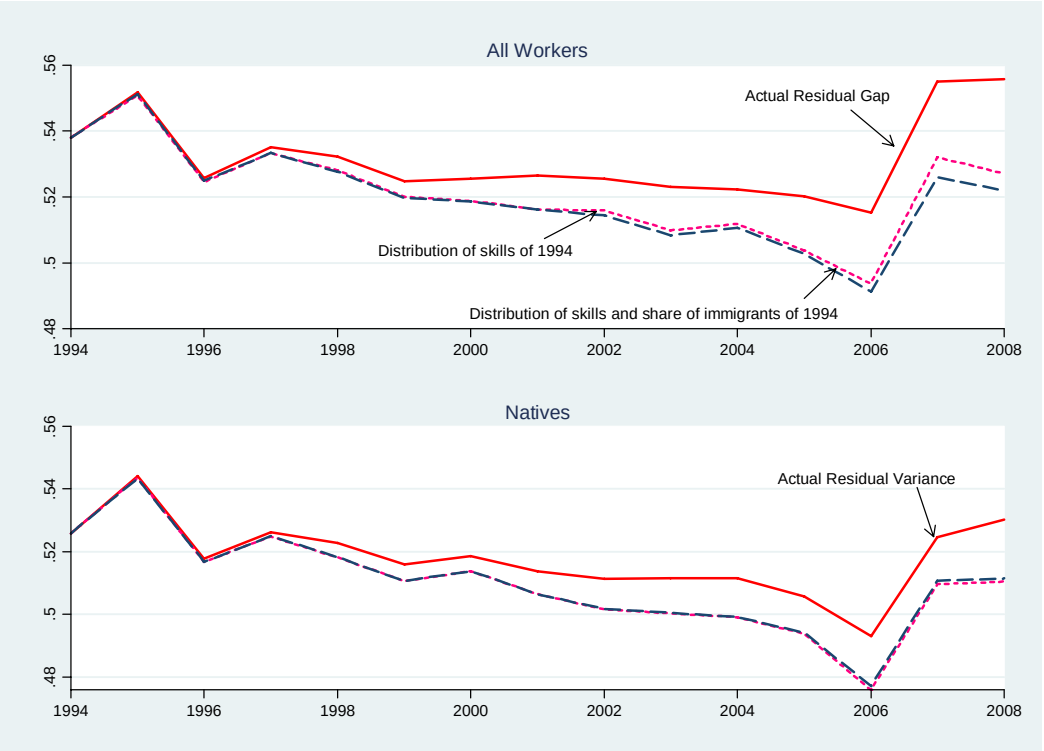
Note: The 90-50 residual gap is based on a standard Mincer wage equation fit separately by year, gender, and immigration status.

Figure 3.10a: 50-10 Residual gap, Men US
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)



Note:. The 90-50 residual gap is based on a standard Mincer wage equation fit separately by year, gender, and immigration status.

Figure 3.10b: 50-10 Residual gap, Men UK
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)



Note: The 50-10 residual gap is based on a standard Mincer wage equation fit separately by year, gender, and immigration status.

Figure 3.10c: 50-10 Residual gap, Women UK
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)

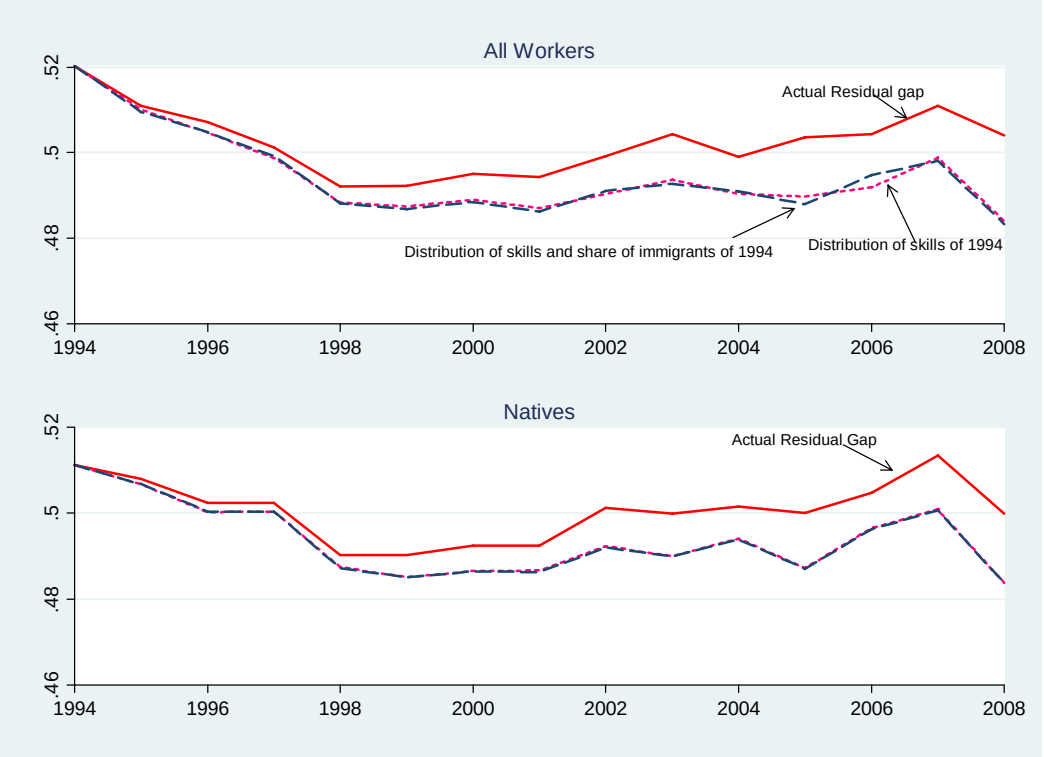
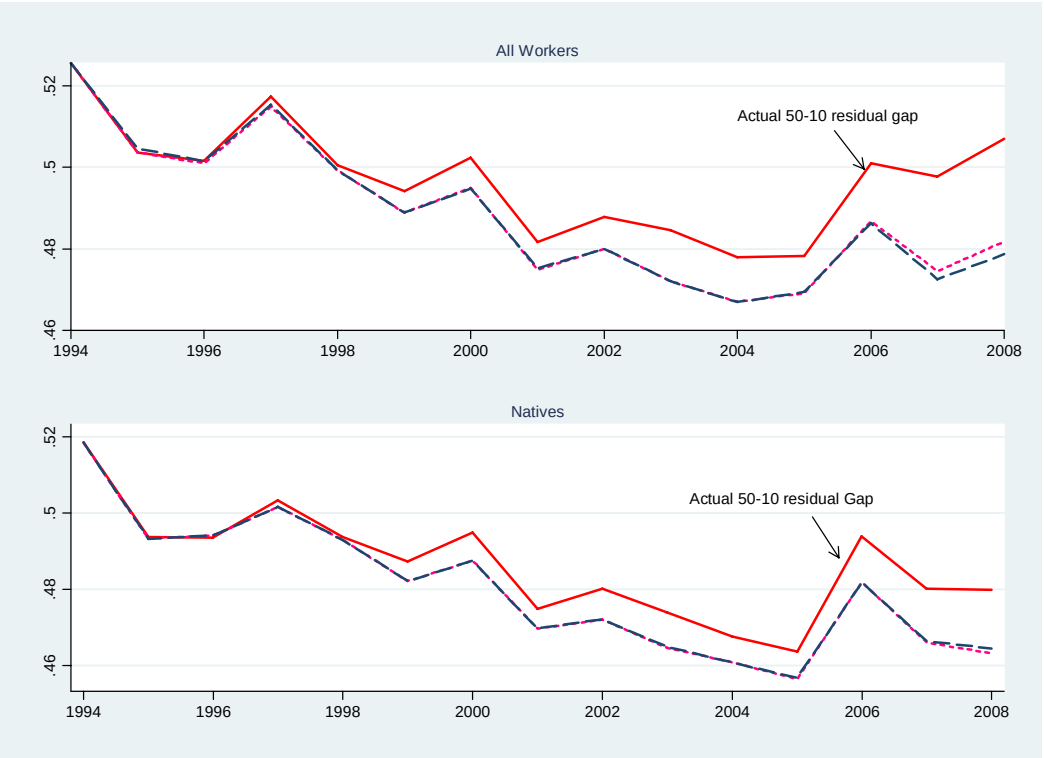


Figure 3.10d: 50-10 Residual gap, Women UK
(Holding Distribution of Skills and Share of Immigrants at 1994 Level)



Appendix A.3.1: Cohort effects

One of the aims of this paper is to account for the unmeasured aspects of human capital. Due to the difficulties of investigating this by following the limited literature (Lemieux 2006, Gould and Moav 2008), this paper uses as a proxy for unobservable skills the residual of a standard Mincer equation; in other words, to interpret the residual as the unmeasured aspect of human capital, I model the specification only on age, educational level (years of schooling) and experience.

As a sensitivity check of my results, I recomputed the results for all workers as well as natives and immigrants separately by adding other controls, such as a dummy variable for area of origin and years spent in the receiving country by the immigrants. As should be expected, the level of the results decreases slightly; nevertheless, the trends do not dramatically change, and more importantly, the wage inequality due to residuals between natives and immigrants remains.[43]

To check whether the trends and levels for immigrants are driven by cohort effects, i.e., cohort of arrivals in the receiving country, Tables A3.1a and A3.1b present the main results by examining the wage inequalities of male immigrants in the US and UK in three different years (1994, 2000 and 2008).

The main measures used are the standard deviation of log hourly wage and the residual variance. However, the analysis does not show any dominant trends of the arrival cohort.[44]

[43] Results for the sensitivity check are available on request from the author.

[44] Considering male immigrants in the US, about 30% of them arrived between 1978 and 1989, followed by 26% who arrived during the 1990s. Only 8% of immigrants working in the US arrived before 1967, and 18% arrived between 1999 and 2008. For the UK, 37% of male immigrants arrived before 1967, while about 15% of them arrived, respectively, during the 1970s, 1980s and 1990s. Another 18% of male immigrants in the UK arrived between 1990 and 2008.

Table A.3.1a: Analysis by Cohort of Arrival, Male Immigrants, US

Cohort of arrivals	1994	2000	2008
Before 1967			
Standard dev.	0.555	0.592	0.635
Residual variance	0.224	0.268	0.299
1968-1977			
Standard dev.	0.565	0.573	0.611
Residual variance	0.205	0.225	0.227
1978-1989			
Standard dev.	0.556	0.550	0.579
Residual variance	0.208	0.209	0.229
1990-1998			
Standard dev.	0.613	0.562	0.599
Residual variance	0.243	0.197	0.215
1999-2008			
Standard dev.		0.571	0.607
Residual variance		0.195	0.223
Stand. dev. all Immigrants	0.599	0.581	0.613
Residual Variance Immigrants	0.227	0.218	0.232

Table A3.1b: Analysis by Cohort of Arrival, Male Immigrants, UK

Cohort of arrivals	1994	2000	2008
Before 1967			
Standard dev.	0.527	0.557	0.544
Residual variance	0.213	0.244	0.217
1968-1977			
Standard dev.	0.625	0.561	0.595
Residual variance	0.270	0.241	0.300
1978-1989			
Standard dev.	0.631	0.615	0.570
Residual variance	0.302	0.281	0.281
1990-1998			
Standard dev.	0.793	0.675	0.611
Residual variance	0.473	0.353	0.275
1999-2008			
Standard dev.		0.631	0.594
Residual variance		0.291	0.281
Stand. dev. Immigrants	0.611	0.602	0.591
Residual Variance Immigrants	0.292	0.285	0.277

Chapter 4: Real Wages, Wage Inequality and the Regional Cost-of-living in the UK

4.1: 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; Schmitt, 1995 and subsequent chapters of this thesis). 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 overall 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 graduate 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, measured as the gross monthly rent, 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 Lemiuex (2001), Gosling and Lemieux (2004), Machin and Van Reenen (2008) 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 an increase in both the 90-50 and 90-10 wage differential (Bell and Van Reenen, 2010); however, as documented in Chapter 3 between 1994 and 2008 the 90-50 log wage gap increased in both the UK and the U.S, though in the same time period women in the UK experienced a slight decrease. Both countries also 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. 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 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. Therefore only London looks to have disproportionately more graduates relative to its population share. 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 16.5% of workers have a high school degree and 13.1% less than high school. 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,[45] 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 explanations 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; Moretti (2010) explains that the increase in the cost-of-living is somehow correlated with the increased attractiveness of those cities.

Moretti (2004) explains that if externalities exist, we should expect that plants located in cities with high levels of human capital to produce a higher output with the same inputs than otherwise similar plants located in cities with low levels of human capital. Specifically he presents a model showing that firms are more productive in cities with high overall levels of human capital, because of spillovers. In equilibrium, firms are indifferent between cities because wages are higher in cities with a higher overall level of human capital, so the unit costs are the same elsewhere. Workers are indifferent because housing prices are higher in cities with a higher overall level of human capital.

In the following model developed by Moretti (2004) there are two cities and two types of labor, educated and uneducated workers. There are two types of goods, a composite good y-nationally traded- and land h- locally traded. Each (city) firm is a competitive economy that produces y using a Cobb-Douglas technology: $y = AH^{\alpha H}L^{\alpha L}K^{\beta}$ where H and L are the hours worked by educated

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

and uneducated workers, respectively and K is capital. To consider how human capital spillover enters the model, the productivity of plants in a city depends on the aggregate level of human capital in the city: A=f(S), where S can be measured as the fraction of college-educated workers in the city, outside the firm. In the absence of human capital spillovers from education, $\delta f/\delta S=0$. In this case, productivity of a firm increases if more skilled workers are employed in the firm, but holding constant the firm's workers in the city have no effect on productivity. If college share in a city generates positive human capital spillovers, a rise in college share raises productivity of all plants in the city.

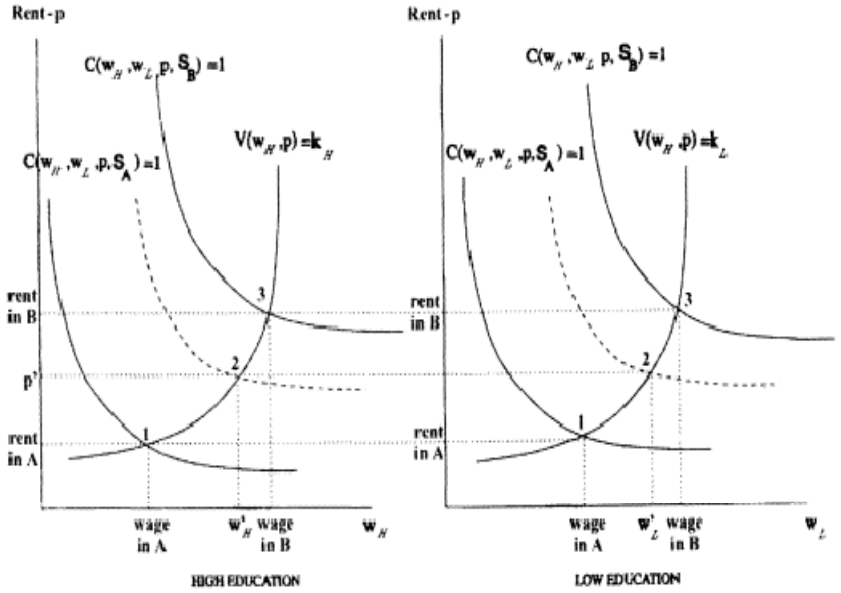
Because the composite good, y, is traded, its price is the same everywhere. Variation in the cost of living depends only on variation in cost of land,p, which is the same for all workers in a city. Workers maximise utility subject to a budget constraint by choosing quantities of the composite good and residential land.

Workers and firms are perfectly mobile, and profits are assumed to be zero. Equilibrium is obtained when the utilities of workers in all cities are equal and firms in different cities have the same unit costs.

Figure 1 taken from Moretti (2004) presents an example of two cities A and B, two education groups (high skilled H and low skilled L) where the indirect utilities, respectively $V_H(w_H, p)$ and $V_L(w_{L,} p)$ of those two groups are a function of nominal wages and cost of land. The indifference curves of the two groups are represented in the figure by the upward-sloping lines in each panel.

The downward-sloping lines show combinations of wages and rents that hold constant firms' unit costs: C (wH, wL, p, r, S)=1; r represents the price of capital, represented by r, is assumed to be constant across cities; S enters the cost function if human capital externalities exist. Firms can produce the same level of output but using less labour and capital in cities with more human capital. Point 1 in the left panel represents the equilirium combination of wage of educated workers and cost of land in city A. The same point (1) in the right panel desplays the same combination for less educated workers. Suppose that due to technological differences, skilled workers are more productive (and more concentrated) and more requested in city B than in city A; city B also pays higher wages to skilled workers, therefore encouraging skilled workers move to B.

In the absence of spillovers the equilibrium in city B is represented by point 2; if spillover exists then the isocost curve shifts further to the right with the magnitude of the spillover represented by the distance from 2 to 3.



Notes: Point 1 is the equilibrium in city A. Point 2 is the equilibrium in city B without externality. Point 3 is the equilibrium in city B with externality. The dashed lines in both panels are the isocost curves in city B without externality. W_H and w_L are the nomonal wages of educated and uneducated workers, respectively.

Bertrand et al.(2008) conclude that industries move towards workers more readily then workers move towards industries; firms reallocate production of skill-intensive industries and products toward regions whose relative factor prices best match their factor needs.

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 n 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 of household'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) calculate 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 points[46]. 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 how different measures of inflation used can lead to differences in 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

[46] 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.

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 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 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 price 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% of incomes. 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 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 how much the cost-of-living accounts for the observed 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 chapter 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.

4.2 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.

4.2.1 National RPI: descriptions and drawbacks

The RPI is defined (ONS, 2007)[47]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[48]; 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.

The data used to produce the weights[49] 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

[47] Consumer Price Indices Technical Manual , page 3, 2007.
[48] Details on the collection of the data in the RPI are presented in the Appendix.
[49] 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.

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-living[50] 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 home-owners 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

[50] 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.

(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 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 mortgagers’ 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 percentage 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 developed

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 % 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.

4.2.2 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 rent[51] from 1997 to 2008, derived from the household data from FRS carried out jointly by the ONS and the National Centre for Social Research[52].

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 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,

[51] 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.
[52] Because data do not exist before 1997, it has not been possible to extend the analysis to earlier time periods.

the measure used is the cost of housing, specifically gross rent faced by households in region r (r=1…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 weights[53] 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.

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

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

(Moretti, 2010). The use of rents as proxy for housing costs is one of the approaches[54] 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 groups[55]. This is motivated by the idea of generating 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, the derived average weekly rents 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

[54]See Diewert (2003) for a survey of alternative approaches.
[55] This means that the measure used by Moretti will possibly add more variation to the final estimates.

as well as the size of housing, due to the small sample the disaggregation by region would lead to imprecise estimates[56].

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 assume that the two regions are characterised by different population size specifically in Region 1 is concentrated 25% of the national labour force and in Region 2 being characterised by a higher concentration of the labour force, that is 75% of the national labour force let us assume that Region 1 has a lower 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 relatively 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

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

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% higher.

This simple example illustrates how the regional deflation on the average national wage may differ in magnitude depending on the combination of other factors. Specifically, the final estimates 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 role of the regional deflation.

Table A: Illustrative example

	Share Graduate	Share High school	Wage Graduate	Wage High school	Region. 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	120.8	
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 its role 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 9, comparing wages deflated by the Regional RPI to the nominal wage reported in parenthesis, the level of the absolute graduate high school wage gap decreases 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 (1)	**Reg.** RPI (2)	**G** Share (3)	**HS** Share (4)	**LHS** Share (5)	**G** Hourly Wage (6)	**HS** Hourly Wage (7)	**LHS** Hourly Wage (8)	**G-HS** Wage Gap (9)	**G-HS** Wage Ratio (10)	**G-LHS** Wage Gap (11)	**G-LHS** Wage Ratio (12)
London	11.7	113.5	22	14.3	7.6	11.22	8.18	6.6	3.04	1.37	4.62	1.7
						(12.69)	(9.27)	(7.50)	(3.42)	(1.37)	(5.19)	(1.69)
South East	14.7	105.5	16	17.2	13.2	12.2	8.46	6.67	3.74	1.44	5.53	1.83
						(12.83)	(8.9)	(7.00)	(3.93)	(1.44)	(5.83)	(1.83)
Rest of the UK	73.6	97.2	62	68.5	79.2	11.43	7.64	6.37	3.78	1.5	5.06	1.79
						(11.16)	(7.47)	(6.22)	(3.69)	(1.49)	(4.94)	(1.79)
National						11.55	7.87	6.43	3.64	1.47	5.12	1.8
						(11.76)	(7.97)	(6.42)	(3.9)	(1.48)	(5.34)	(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.

4.3. 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 dummy[57] .

As pointed out by Manacorda et al. (2007) in analysis of UK data it is standard practice to define education by the highest level of qualifications obtained. This criterion is particularly appropriate in a cross-country comparison due to the difference in the schooling systems, and in fact the highest level of education has been the criteria adopted in chapter 2 and chapter 3.

Because this chapter only focuses on the UK and extends the analysis to migrants, using the highest level of education is not longer appropriate since the LFS does not always record the education of immigrants accurately. In fact a large proportion of migrants in the LFS reports holding "other qualification" suggesting that the foreign qualifications are classified in this category; on the other hand for the UK born workers the response of "other" category is likely to refer to a very low level of education. Therefore to overcome these limitations

[57] Following Wooldridge (2006) suppose to deflate wages of 2008 at 1997 denoting the deflator factor for 2008 wages P2008 (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 ($wage_i$/P2008) =log($wage_i$)- log(P2008).
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.

this chapter I classify education by using the variable "age left full time education". 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" therefore all universities graduates will be included in the latter category even if they have 15 years of schooling. 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:

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

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,…..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:

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

Additional specifications also add region fixed effects to the baseline regressions[58].

[58] The same net effect (on the graduate coefficient) to the regional deflation could be obtained by adding region-year dummy interaction terms (since deflating is common to one region in any one year), however the approach adopted in this work makes clear the effect on levels.

4.4. 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 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.[59]

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

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

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 postcode. The LFS also contains information at regional level, where region is determined according to usual residence. The LFS identifies 20 regions[60]. 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

[60] 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.

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, 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.[61]

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

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

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.

4.5. Results

4.5.1 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.

Figure 4.1 together with Table 4.1 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 4.1. 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 4.2 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 4.3a, 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 4.2 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 4.3a 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 4.1 and Table 4.1 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.

The differences in the cost of housing faced by workers of different levels of education could also reflect differential changes in quality of housing (Moretti, 2010).

However as explained later on in this chapter (page 166 and page 167) a very well know bias in the construction of any cost-of-living index is that the "constant basket" approach ignores the quality in existing goods and services and this is mainly due to the fact that quantity data are usually unavailable to the researcher.

To show that not all variation in house prices across regions can be explained by differences in quality, table 4.3b based on the Survey of English Housing for the year 2005, use as proxies for quality of housing across the UK regions the availability of a spare room, the kitchen wide over 6.5ft and the availability of a garden. The SEH does not contain any information on Scotland,

Wales, and Northern Ireland therefore Table 4.3b only refers to the 9 English regions.

Considering the availability of a spare room reported in percentage in column 1, houses in London displays the lowest percentage (39%) while almost 63% of the houses in the North East have a spare room; similarly when looking at the measures of kitchen, column 2 reports the percentage of housing having a kitchen wide more than 6.5ft; again houses in London, though being more expensive, have a lower percentage of kitchen wider than 6.5 ft. Considering the percentage of houses having a garden, 82% of the houses in London have one while expect the Yorkshire and the Humber, in the remaining regions more than 90% of the houses have a garden. The statistics reported in columns 1-3 demonstrate that even though the cost of housing are higher in London, they do not necessarily reflect a higher quality of housing.

To better understand if and by how much the national and regional RPI diverge, Figure 4.3 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 4.4 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 4.5 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 East[62]; 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 4.6 reports some common measures of wage inequality based on the log hourly 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 90^{th} and 50^{th} 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 4.6 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

[62] 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.

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 4.5, 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 4.7 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 4.8 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 4.9 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 rows of table 4.9 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 4.4 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 4.10 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 4.11 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 4.5a

and 4.5b 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 4.5b), though the gap is lower.

Figures 4.6a and 4.6b 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 4.6a 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 4.6b 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 4.7 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 4.12 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).

4.5.2 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 4.11 are due to regional differences in the cost of housing in the UK. Table 4.13 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 an 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 deflator[63]. However this fall

[63] 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

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.

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

---

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.

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 4.12, 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 abilities[64]. 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 4.13 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, Manacorda et al. (2007) provide evidence that the native-immigrants wage differential is sensitive to the share of immigrants in the working age population; in fact they

[64] 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.

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 4.14 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 4.15 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 4.16 reports similar estimates to tables 4.14 and 4.15 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 4.16 reports results from similar models to those in table 4.14. 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: table 4.3A of the appendix shows how the tenure of workers changes by education and age: the share of workers aged 35 or less renting is higher than when considering the whole sample (Table 4.2A). 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 4.17 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 4.17 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.

4.6. 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 chapter 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).

Appendix 4.1

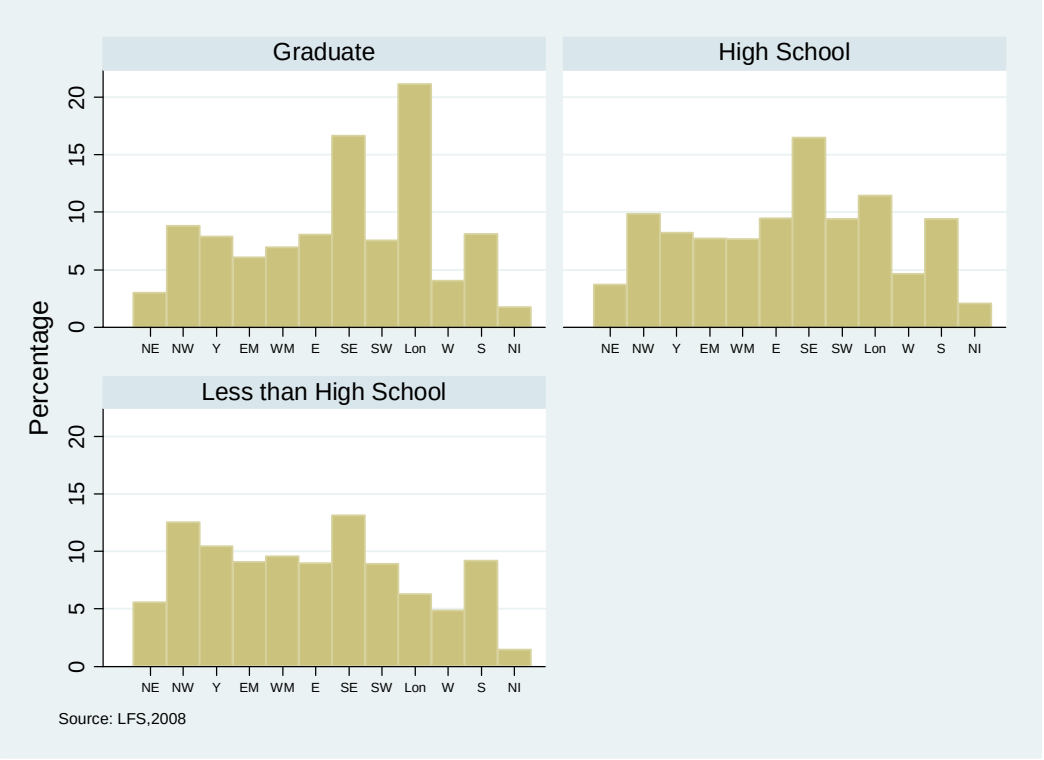
A4.1: The RPI construction

RPI price data are collected from a range of shopping locations across the UK. Locations are defined as clusters of enumeration districts broadly representing a central shopping area and the areas where the local shopping population tend to live. The sample is stratified by region (Scotland, Wales, Northern Ireland and the nine English Regions) and size of location. The number of locations per region is proportional to aggregate regional expenditure by households. In total, there are around 150 locations, with the number of locations per region ranging from 5 in Northern Ireland to 22 in the South East. In each location, a random selection of outlets is chosen from a sampling frame based on an enumeration by price collectors of all the outlets in the location. Prices are collected from the randomly selected outlets for over 550 pre-defined items. The latter are purposively selected to represent typical expenditure in each of the 85 different categories of expenditure, called "sections". For most items, it is sufficient for the purposes of calculating the national RPI to collect one price quote only in each location. This means that the number of quotes per region for most items is similar to the number of locations per region. Each month over 110,000 price quotes are collected locally but the number of regions varies from around 4,000 in Northern Ireland to over 16,000 in the South East, reflecting the sizes of the local shopping population and the amount they spend. Each month over 10,000 prices are also collected centrally from major chains with national pricing policies; these are combined with the locally collected prices according to the proportion of expenditure they represent locally; this is where ONS gathers prices direct from retailers to construct centrally calculated indices. Although prices may vary by region, there is no regional dimension to the calculation itself.

Tables A.4.2a and A.4.2b report respectively the average rent per person and the share of housing expenditure by region from 1997 and 2009. As documented in table A.4.2a, the average rent per person has been increasing over the sample period in all regions; however London, the South East and the East have the highest rent. Table A.4.2b reports the share of housing expenditure based on the FES from 1997 and 2008; London displays the highest share overall in the UK; in 2008 housing costs made up 29% of the total consumption for Londoners; Northern Ireland presents the lowest share for housing, with 14% in 2009. Table

A.4.2c reports the regional RPI for all regions over time from 1997 and 2009. The table provides evidence of how much the regional PRI differ by the national RPI. Once again London displays a higher regional RPI over time and within regions and higher than the national RPI; this implies that the national RPI provided by the ONS does not properly reflect the cost of living faced by workers in the London area. Similarly, in some regions, the cost of living experienced is lower than the one painted by the national RPI (Northern Ireland, North East and Yorkshire). Table A.4.2d reports the percentage of workers with a graduate degree; in 2008 about 43% of workers in London had a graduate degree; followed by the South East with about 27% of graduates, the share in the remaining regions falls up to 17% in the North East. Figures A.4.2a and A.4.2b plot the characteristics of tenure by level of education in all years; while figure A.4.2a presents the share for all workers, figure A.4.2b reports the characteristics of tenure for younger cohorts, those aged 35 or less. It is quite clear that younger cohorts are more likely to rent rather than paying mortgages or owning their own house. However, the fact that on average the great majority of workers pay mortgages recalls the need to calculate three different indices per tenure (rent, owner, mortgages) pointed out by Crawford and Smith (2002) and based on the motivation according to which a change in prices of goods or services (e.g. housing) can only affect the measure of inflation if the household pays for those goods or services.

Figure 4.1: Distribution of Workers by Education Within Region, All Workers, 2008



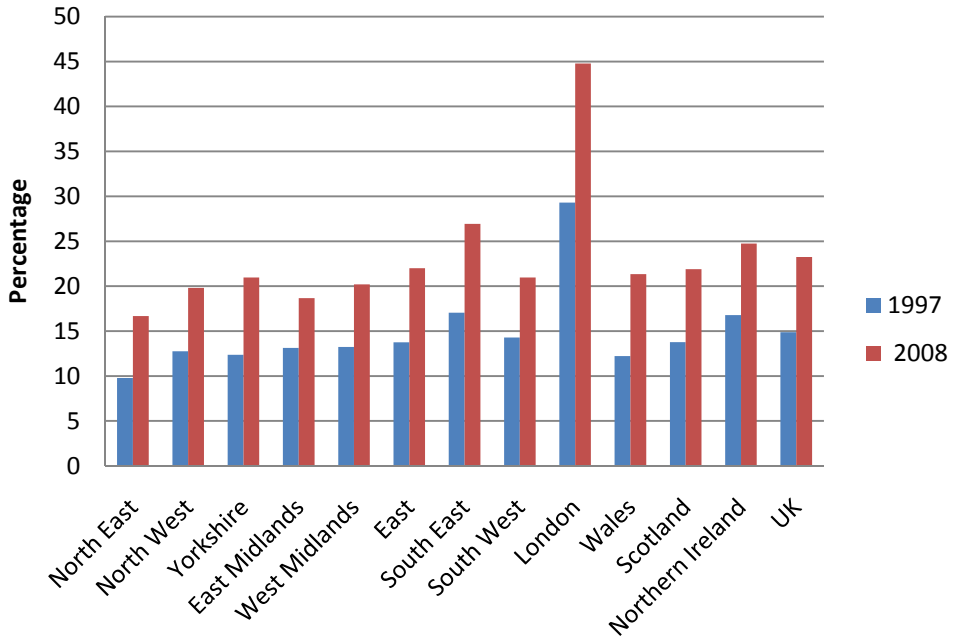
Notes: NE= North East; NW=North West; Y=Yorkshire; EM=East Midlands; WM=West Midlands; E=East; SE=South East; SW=South West; Lon=London; W=Wales; S=Scotland; NI=Northern Ireland. Sample is based on men (16-64) and women (16-59) working full time and part time, employees and main job only. Graduate refers 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 4.1: 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 4.2: 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 4.2: 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 4.3a: Rental Prices 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 4.3b: Variation in quality measure of house, 2005.

Regions	Variation in House Characteristics		
	Has spare room	Kitchen over 6.5ft wide	Has a garden
	(1)	(2)	(3)
North East	62.7	93.9	86.3
North West	56.4	93.6	86.0
Yorkshire and the Humber	58.1	88.6	92.3
East Midlands	58.2	92.4	96.2
West Midlands	59.6	92.6	96.1
East	62.0	92.4	95.2
London	39.1	85.3	82.2
South East	57.6	92.8	94.5
South West	60.4	91.8	93.1

Notes: Based on Survey of English Housing 2005.

Table 4.4: 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 4.5: Real Hourly Wage by Aggregate Regions, 1997-2008

	UK		London		South East		Rest of the UK	
Year	**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 4.6: 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 4.7: 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 4.8: Changes in Real Hourly Wage by Region, 1997-2008

Region	Based on National RPI	Based on Regional RPI
	(1)	(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 4.9: Changes in Weekly Rent, by Education Group

		1997	
	UK	**London**	**South East**
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%

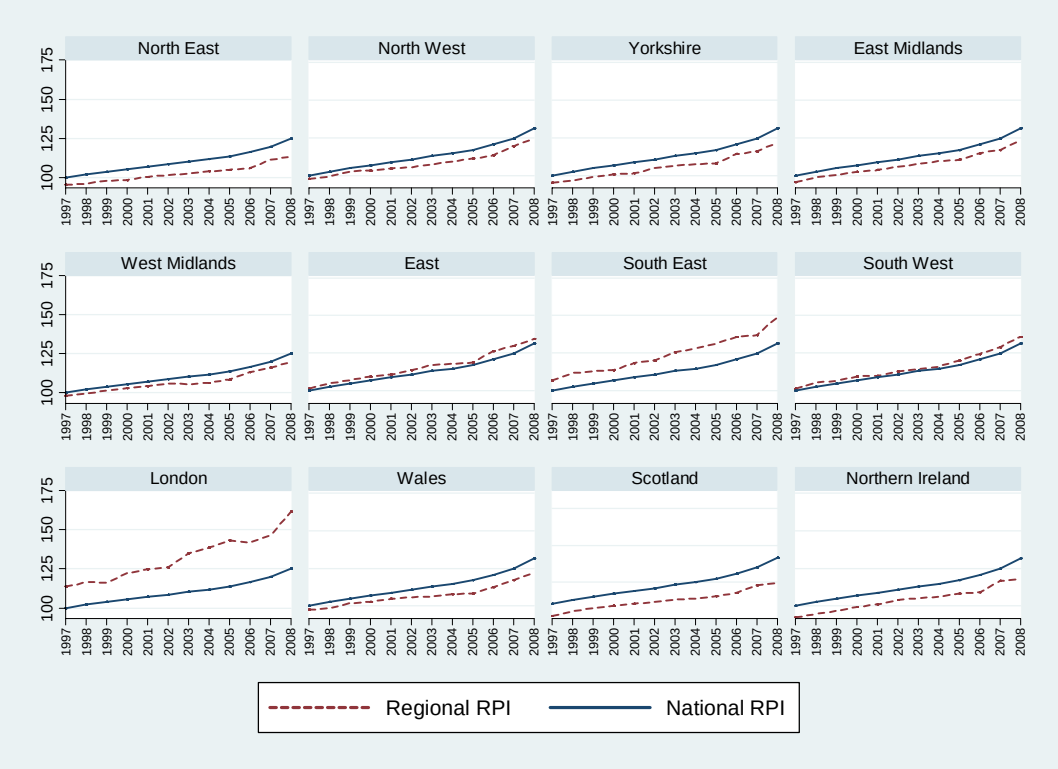
Based on the FRS.

Table 4.10: 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%

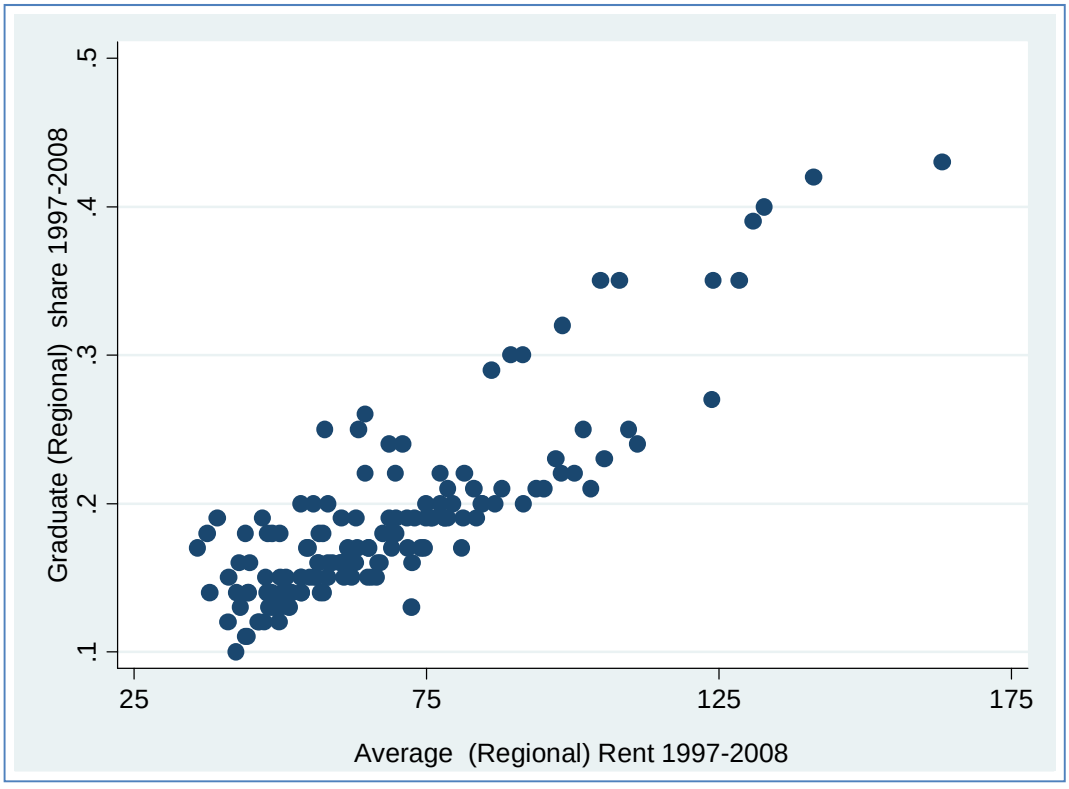
Notes: Based on the LFS. Sample is based on men (16-64) and women (16-59) working full-time and part time, employees and main job only. Wages are deflated at 1997 level. Graduate refers to anyone who left full time education at 21 or later; "High school graduate" refers to anyone who left full time education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

Figure 4.3: National and Regional RPI, UK 1997-2008



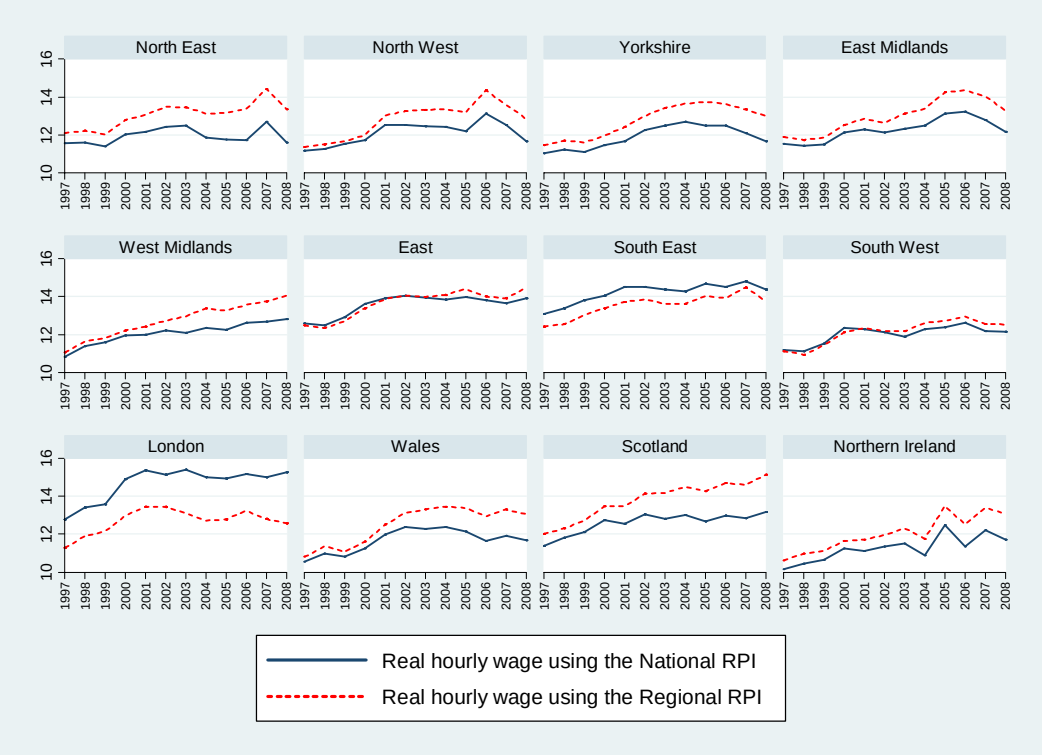
Source:EFS/FES/LCS and FRS

Figure 4.4: How Increasing Share of Graduates Relate to Increasing Price for Rent



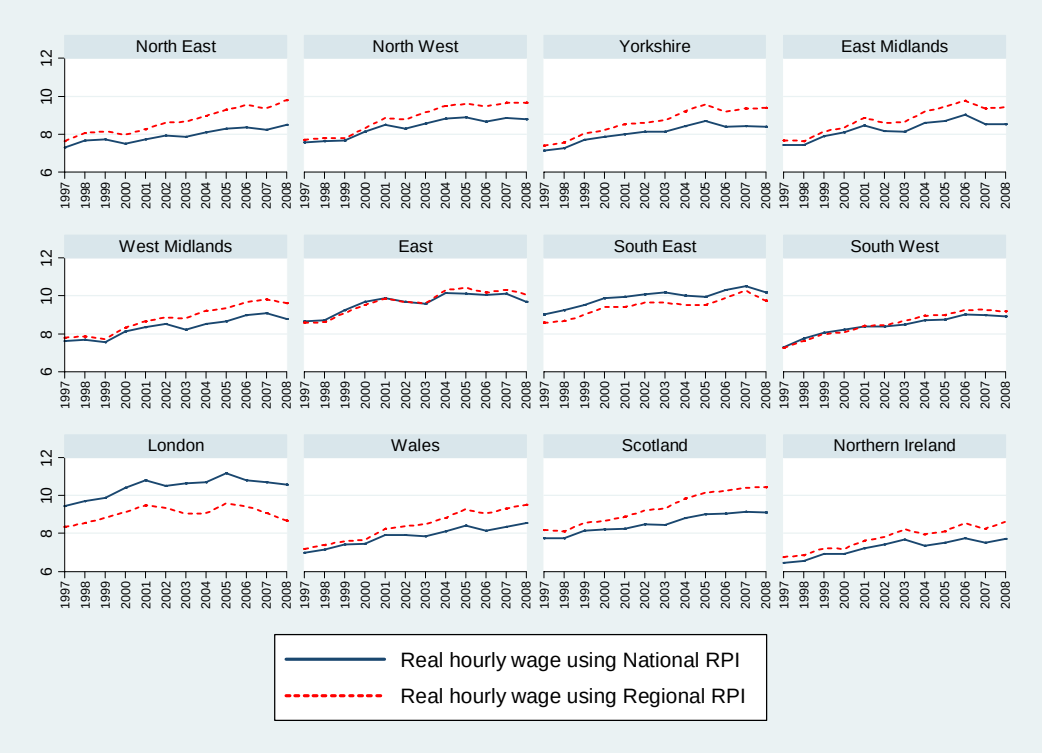
Notes: Based on the FRS and LFS, 1997-2008.

Figure 4.5a:Real Hourly Wage for Graduate Workers by Region, All Workers 1997-2008



Notes: Based on the LFS. Graduate refers to anyone who left full time education at 21 or later.

Figure 4.5b: Real Hourly Wage for High School Workers by Region, All Workers 1997-2008



Notes: Based on LFS. High School workers are defined as those who left school between 17 and 20 years old.

Table 4.11: Real Hourly Wage for Graduate Workers by Region, 2008

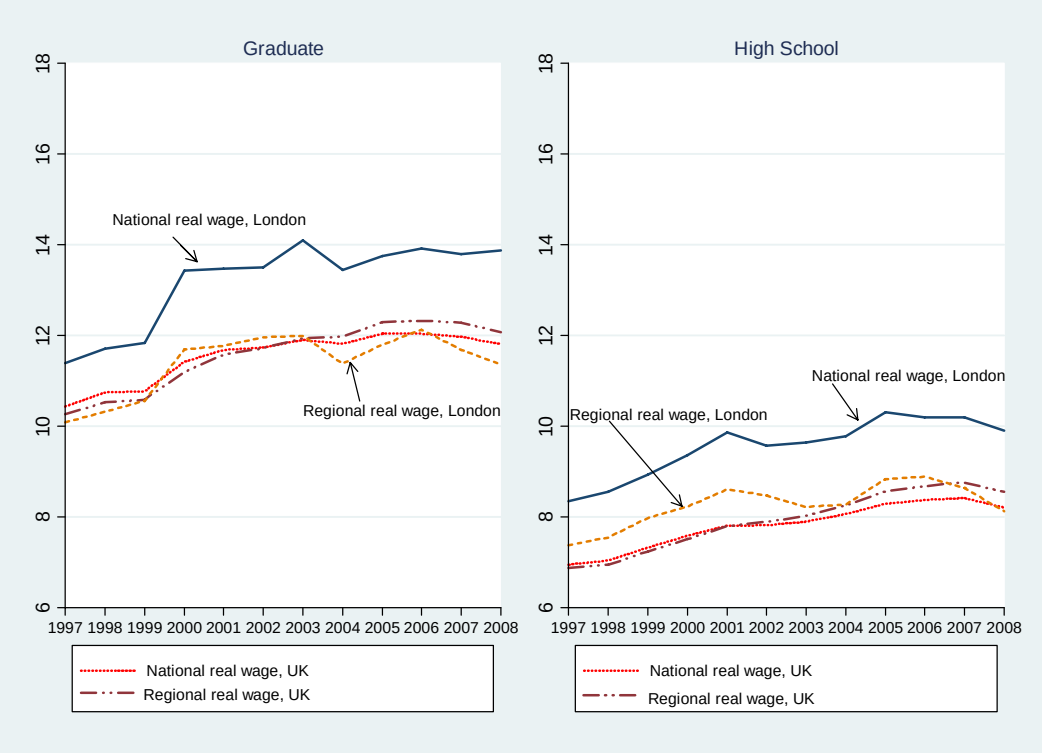
Region	based on National RPI	based on Regional RPI	Difference
	(1)	(2)	(3)
North East	11.23	12.95	1.72
North West	11.32	12.46	1.14
Yorkshire	11.27	12.59	1.32
East Midlands	11.79	13.01	1.22
West Midlands	12.43	13.61	1.18
East	13.27	13.80	0.53
South East	13.83	13.23	-0.60
South West	11.79	12.14	0.35
London	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 4.6a: Real Hourly Wage for Graduate and High School Male Workers, London and UK 1997-2008

Notes: Based on the LFS. 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 4.6b: Real Hourly Wage for Graduate and High School Female Workers, London and UK 1997-2008



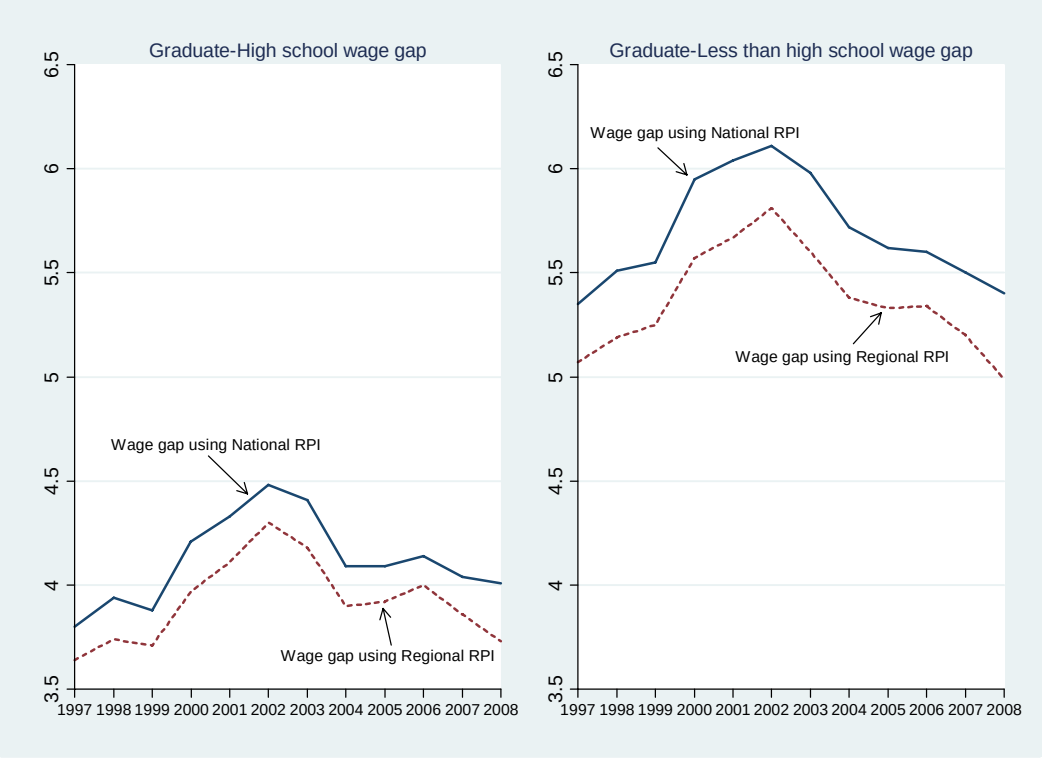
Notes: Based on LFS. Graduate workers are defined as those who left school at age 21 or more. High school workers are defined as those who left school between 17 and 20 years old.

Table 4.12: 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	Diff.	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. Graduate refers to anyone who left full time education at 21 or later.

Figure 4.7: Graduate-High School and Graduate-Less then High School Wage Gap, UK 1997-2008



Notes: Based on the LFS. 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 4.13: 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)
Model 1						
Nominal Wage Dif.	0.419***	0.402***	-0.017	0.408***	0.388***	-0.014
	(0.01)	(0.017)	(0.00)	(0.008)	(0.007)	(0.00)
Model 2						
Real Wage Dif.	0.413***	0.395***	-0.018	0.410***	0.382***	-0.028
(Regional RPI)	(0.009)	(0.016)	(0.00)	(0.007)	(0.012)	(0.00)
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 4.14: 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	1997	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)
Model 1						
Nominal Wage	0.408***	0.383***	-0.025	0.399***	0.402***	0.003
Difference	(0.01)	(0.014)	(0.00)	(0.012)	(0.011)	(0.00)
Model 2						
Real Wage dif.	0.401***	0.363***	-0.038	0.371***	0.364***	-0.007
Regional RPI	(0.007)	(0.01)	(0.00)	(0.011)	(0.011)	(0.00)
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 4.15: Nominal and Real Conditional Wage between Workers and with College or more and High School Degree, by Men and Women

	1997	2008	1997-2008 Change	1987	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Men						
Nominal Wage Diff.	0.366***	0.370***	0.004	0.357***	0.354***	-0.003
	(0.009)	(-0.016)	(0.00)	(0.006)	(0.01)	(0.00)
Real Wage Diff.	0.360***	0.351***	-0.009	0.359***	0.348***	-0.011
Regional RPI	(0.007)	(0.009)	(0.00)	(0.006)	(0.01)	(0.00)
Panel B. Women						
Nominal Wage	0.466***	0.427***	-0.029	0.452***	0.413***	-0.039
Difference	(0.015)	(0.019)	(0.00)	(0.015)	(0.012)	(0.00)
Real Wage Diff.	0.460***	0.411***	-0.049	0.454***	0.407***	-0.047
Regional RPI	(0.014)	(0.015)	(0.00)	(0.014)	(0.015)	(0.00)
Region Fixed Effects	No	No		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 4.16: Additional Specifications: Nominal and Real Conditional Wage between Workers and with College or more and High School Degree: All Working Force Population

	1997	2008	1997-2008 Change	1997	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All Workers						
Nominal Wage Diff.	0.430***	0.397***	-0.033	0.419***	0.383***	-0.036
Difference	(0.009)	(0.016)	(0.00)	(0.011)	(0.011)	(0.00)
Real Wage Diff.	0.421***	0.376***	-0.045	0.422***	0.376***	-0.046
Regional RPI	(0.011)	(0.009)	(0.00)	(0.01)	(0.01)	(0.00)
Panel B. Men						
Nominal Wage Diff.	0.389***	0.380***	-0.009	0.381***	0.365***	-0.016
	(0.007)	(0.02)	(0.00)	(0.008)	(0.014)	(0.00)
Real Wage Diff.	0.381***	0.356***	-0.025	0.384***	0.357***	-0.027
Regional RPI	(0.008)	(0.009)	(0.00)	(0.008)	(0.01)	(0.00)
Panel C. Women						
Nominal Wage Diff.	0.461***	0.406***	-0.055	0.449***	0.393***	-0.056
	(0.015)	-0.015	(0.00)	(0.018)	(0.012)	(0.00)
Real Wage Diff.	0.453***	0.388***	-0.065	0.451***	0.386***	-0.065
Regional RPI	(0.017)	(0.012)	(0.00)	(0.017)	(0.013)	(0.00)
Region Fixed Effects	No	No		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 4.17: 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	2008	1997-2008 Change	1997	2008	1997-2008 Change
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All workers						
Model 1						
Nominal Wage Diff.	0.696***	0.601***	-0.095	0.669***	0.575***	-0.094
	(0.008)	(0.016)	(0.00)	(0.016)	(0.012)	(0.00)
Model 2						
Real Wage Diff.	0.675***	0.562***	-1.113	0.677***	0.562***	-0.115
Regional RPI	(0.013)	(0.017)	(0.00)	(0.013)	(0.017)	(0.00)
Panel B. Men						
Model 1						
Nominal Wage Diff.	0.650***	0.580***	-0.07	0.623***	0.551***	-0.072
	(0.014)	(0.018)	(0.00)	(0.017)	(0.013)	(0.00)
Model 2						
Real Wage Diff.	0.628***	0.536***	-0.092	0.630***	0.537***	-0.093
Regional RPI	(0.018)	(0.018)	(0.00)	(0.015)	(0.015)	(0.00)
Panel C. Women						
Model 1						
Nominal Wage Diff.	0.746***	0.618***	-0.128	0.720***	0.595***	-0.125
	(0.013)	(0.018)	(0.00)	(0.022)	(0.015)	(0.00)
Model 2						
Real Wage Diff.	0.727***	0.584***	-0.143	0.726***	0.584***	-0.142
Regional RPI	(0.017)	(0.017)	(0.00)	(0.019)	(0.017)	(0.00)
Region Fixed Effects	No	No		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.

Appendix 4.2

Table A.4.2a: Price for Rent by Region, 1997-2008 UK

Region	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
North East	42.5	41.1	44.1	44.4	47.3	49.8	52.0	55.8	58.1	61.6	65.2	71.8
North West	48.9	49.9	53.6	53.6	56.6	56.5	60.7	65.0	69.5	71.7	78.1	83.5
York& Humber	43.2	42.5	44.5	47.7	50.0	55.1	58.2	58.9	65.0	72.4	73.0	78.6
East Midlands	44.6	48.0	47.5	51.0	53.6	58.0	61.7	65.6	67.6	74.9	75.7	81.3
West Midlands	48.5	49.5	51.5	53.7	57.5	62.1	60.4	63.1	68.7	77.4	79.6	84.4
East	56.9	60.9	62.8	66.5	69.1	74.0	81.0	81.5	83.1	95.1	97.1	100.4
South East	69.9	78.1	74.9	76.1	86.7	88.0	98.1	101.8	105.4	111.2	109.6	123.9
South West	56.3	62.2	60.3	66.7	67.2	72.5	74.6	78.6	84.5	91.6	93.7	103.2
London	86.1	91.5	89.5	98.3	104.8	108.0	124.1	128.5	130.8	132.8	141.2	163.2
Wales	46.4	46.3	50.4	51.3	54.8	56.7	57.3	60.8	63.0	68.6	69.8	77.4
Scotland	38.0	41.2	43.0	44.9	47.9	48.6	49.9	54.6	58.2	60.4	64.5	69.7
Northern Ireland	35.9	37.5	39.2	44.1	47.0	53.6	55.8	57.6	63.4	64.5	70.9	68.6
UK	51.4	54.1	55.1	58.2	61.9	65.2	69.5	72.7	76.4	81.9	84.9	92.2

Notes: These data relate to rents derived from the Family Resources Survey.

A.4.2b Share of Housing Expenditure by Region, 1997-2008 of housing expenditure by region, 1997-2008

Region	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
North East	0.22	0.20	0.24	0.24	0.20	0.21	0.22	0.22	0.22	0.25	0.24	0.27
North West	0.19	0.19	0.20	0.22	0.22	0.20	0.22	0.22	0.25	0.25	0.25	0.27
York&Humber	0.20	0.18	0.20	0.20	0.22	0.20	0.20	0.20	0.26	0.22	0.27	0.25
East Midlands	0.20	0.19	0.20	0.20	0.19	0.21	0.21	0.22	0.23	0.24	0.27	0.25
West Midlands	0.21	0.19	0.20	0.21	0.21	0.22	0.22	0.23	0.24	0.24	0.26	0.26
East	0.20	0.20	0.22	0.21	0.22	0.22	0.22	0.23	0.25	0.25	0.27	0.27
South East	0.21	0.20	0.23	0.22	0.22	0.23	0.24	0.24	0.26	0.27	0.27	0.29
South West	0.21	0.22	0.23	0.20	0.23	0.22	0.23	0.23	0.25	0.26	0.28	0.27
London	0.25	0.25	0.24	0.26	0.26	0.26	0.28	0.30	0.31	0.31	0.34	0.34
Wales	0.19	0.19	0.20	0.22	0.20	0.21	0.22	0.23	0.24	0.25	0.25	0.24
Scotland	0.20	0.19	0.21	0.22	0.22	0.22	0.20	0.23	0.23	0.23	0.24	0.27
Northern Ireland	0.16	0.15	0.16	0.16	0.16	0.16	0.16	0.17	0.17	0.20	0.19	0.18

Based on the EFS/FES/LCS

Table A.4.2c Regional and National RPI, 1997-2008

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
North East	95.6	96.3	98.0	98.6	100.6	101.7	102.5	104.0	105.2	106.0	111.3	113.4
North West	98.3	99.6	102.3	102.8	103.8	104.5	105.9	107.6	109.0	110.7	115.7	119.5
York& Humber	96.2	97.4	99.3	100.7	101.2	103.9	105.4	106.1	106.5	111.4	113.0	117.4
East Midlands	96.7	99.1	100.3	102.1	102.9	104.7	106.2	107.8	108.3	112.0	113.7	118.6
West Midlands	98.0	99.5	101.5	102.8	104.2	106.0	105.5	106.4	108.5	113.1	116.1	119.6
East	101.2	103.8	105.7	107.6	108.6	110.8	113.7	114.2	115.0	121.1	123.9	127.4
South East	105.6	109.4	110.3	111.0	114.8	116.0	120.5	122.5	125.2	128.6	129.6	139.2
South West	101.0	104.4	105.2	107.6	108.0	110.2	111.3	112.8	116.1	119.6	123.3	128.8
London	113.6	116.5	116.4	122.2	124.7	126.1	134.9	138.5	143.2	141.9	146.2	161.3
Wales	98.0	98.7	101.5	102.2	103.8	104.7	105.0	106.2	106.7	109.9	113.7	117.6
Scotland	93.4	96.0	97.8	98.9	100.1	101.0	102.3	102.9	104.1	106.1	110.1	111.5
Northern Ireland	93.9	95.8	97.3	99.5	100.9	103.2	104.2	104.8	106.7	107.2	113.2	114.1
National RPI	100.0	102.0	103.7	105.3	108.1	110.6	113.5	115.8	118.4	122.0	125.9	131.5

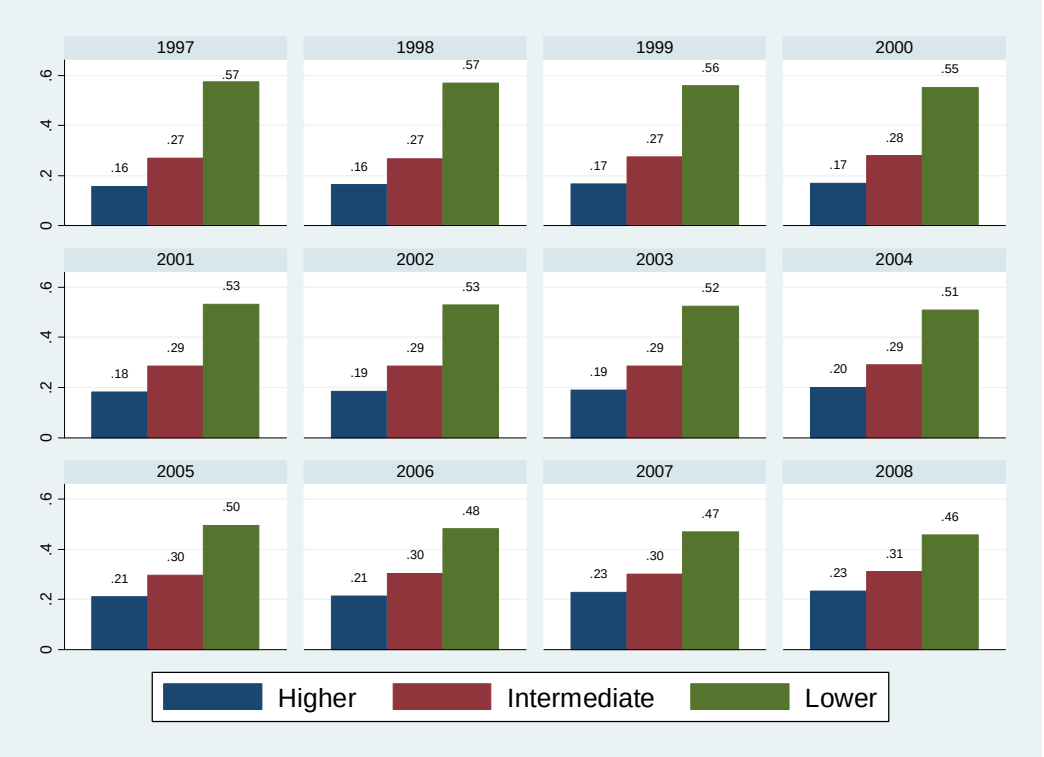
Source: EFS/FES/LCS , FRS and ONS.

Table A.4.2d:Percentage of Graduate Workers, 1997-2008

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
North East	10.3	11.5	10.7	10.6	12.2	12.3	13.6	14.9	14.5	16.6	17.2	16.6
North west	13.1	13.2	14.1	15.3	15.7	15.6	15.8	15.1	17.8	18.9	19.1	19.4
Yorkshire	12.7	14.4	13.9	13.5	15.0	15.2	16.0	16.3	16.7	13.3	18.9	20.8
East Midlands	13.5	13.2	14.5	14.8	15.1	15.1	15.8	14.6	17.8	19.6	19.0	18.7
West Midlands	13.6	13.0	12.7	14.3	14.3	15.5	15.7	17.0	17.5	19.7	20.0	20.1
East	14.1	14.6	16.1	14.8	16.6	16.6	17.4	22.4	20.6	21.0	22.9	21.5
South East	17.5	18.9	19.0	18.8	20.1	21.0	22.2	24.7	23.3	24.4	25.3	26.7
South West	14.6	14.8	15.7	16.4	16.3	15.8	17.4	18.6	19.5	20.4	21.2	20.8
London	28.9	29.8	30.3	31.8	34.6	35.4	34.8	34.9	38.6	39.7	42.3	42.9
Wales	12.2	12.4	13.5	13.9	16.8	17.7	17.7	16.0	19.0	18.5	19.4	21.5
Scotland	14.0	15.1	15.5	15.5	17.6	17.7	17.5	17.1	19.8	18.8	21.9	21.7
Northern Ireland	17.3	18.2	18.7	18.1	19.4	20.4	20.0	25.2	25.2	25.8	23.5	23.8

Notes: Based on the LFS. Sample includes workers aged 16-64,working full time and part time, employed and main job only. Graduate workers are defined as those who felt school at age 21 or more

Figure A.4.2a: Proportion of Workers by Education, 1997-2008,
(Using Age Left School)

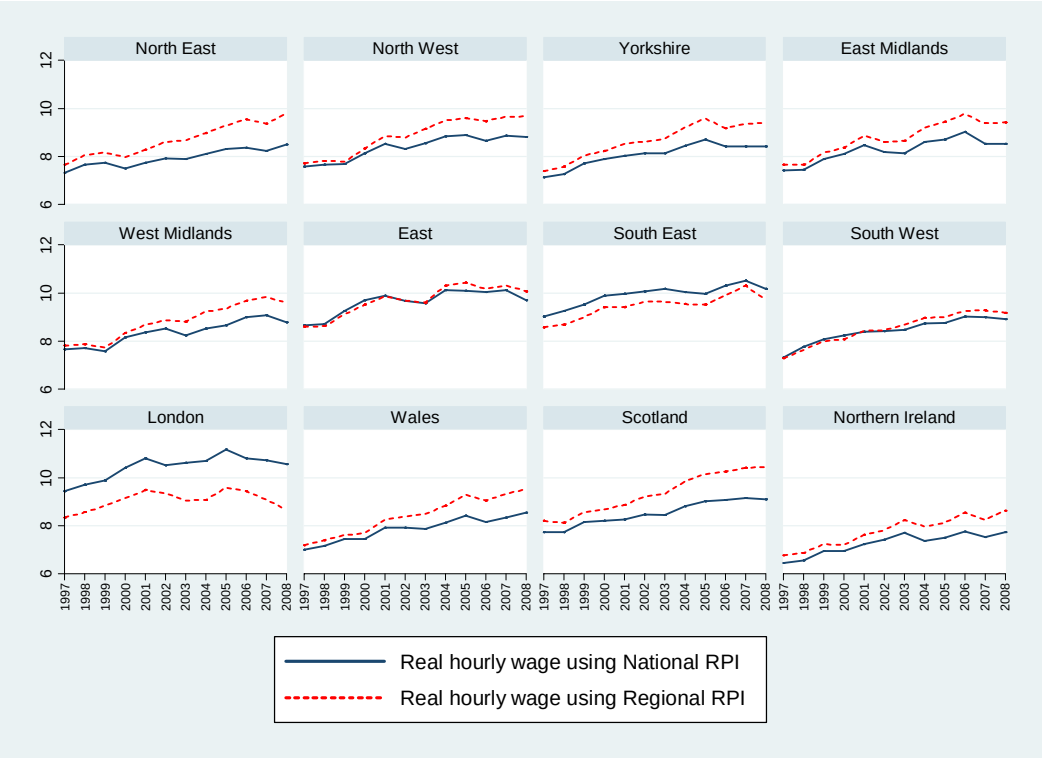


Notes: Based on the LFS. "Higher" workers are defined as those who left school at age 21 or more. "Intermediate" are defined as those who left school between age 17 and 20; "lower" are defined as those who left school at 16 or less.

Figure A.4.2b: Tenure by Education all Workers, London Area 1997-2008

Notes: Based on the LFS. "Higher" workers are defined as those who felt school at age 21 or more. "Intermediate" are defined as those who left school between age 17 and 20; "lower" are defined as those who left school at 16 or less.

Table A.4.2c: National and Regional RPI Using Alternative Price for Housing



Notes: Price for rents are based on the private rents available from the Department for Communities and Local Government.

Chapter 5: Conclusion

The level of wage inequality generated by a country's labour market is of fundamental importance for understanding poverty and social stratification. Despite the existence of a large economic literature investigating causes and trends in wage inequality mainly in the UK and the US, the persistence and increase in wage dispersion in these and other countries still calls for analysis and further research.

This thesis has attempted to investigate the role of composition, immigration and cost-of-living on the increasing wage inequality for the UK and other countries since the late 1980s to 2008.

The empirical analysis of this thesis began by examining to which extent changing characteristics of the labour force can help to account for the fact that residual or within-group wage inequality –wage dispersion among workers with the same education and experience- accounts for most of the growth in wage inequality. The investigation is performed separately for men and women in the UK, US and Italy from 1987 to 2004. This comparative research is motivated by the different experiences of the three countries in terms of wage dispersion, the changing characteristics of the labour force in all countries, as well as the different labour market institutions that could also affect differently levels and trends of wage dispersion (Blau and Kahn, 1996).

The descriptive statistics of chapter 2 has documented that inequality has risen in a number of dimensions in the U.S., the UK and Italy since the late 1980s. Wage dispersion has widened within narrow segments of society and the economy as well. It is not just the share of income going to the top fifth that has risen, but also wage differentials of workers in similar education and experience groups - has exhibited a similar pattern. Not only have more highly skilled and educated workers moved ahead faster than those with less skills and education, but even within the upper groups, those with most skills, and the best education, have moved ahead faster still.

When comparing trends and levels of the increasing residual wage inequality, my cross-country findings present some support for conclusions concerning residual inequality trends proposed by the main empirical literature (Juhn, Murphy, Pierce 1993, Autor, Katz and Kearney 2005, 2008; Acemoglu

2002, Lemieux 2006), however the analysis also reveals some remarkable differences among the three countries, both in the level and sometimes in the direction of the trends of wage inequality. Residual wage inequality accounts for most of the variation in wage inequality from 1987 to 2004, not only for the U.S. but also for Italy and the UK. In the U.S. from 1987 to 2003, changes in the residual for men accounted for the 86% of the overall change in the total variance, and 53% for woman. In the UK, despite the decrease in overall variance of wage both for men (0.033) and women (0.034), the residual variance increased by 0.014 log points. More strikingly, the within-group variance in Italy explains almost all wage inequality, with the bulk of the increase occurring during the 1990s.

When looking for particular explanations for these different trends a single reason is not forthcoming and the composition effects hypothesis appears to hold only for the U.S.; in the UK and Italy when controlling for distribution of skills, both at 1987 and at 2004, the residual variance still explains the bulk of the increase, suggesting a more marginal role for the composition effects. In partial contrast with AAK (2005) who show that composition effect was concentrated in the lower tail of the earnings distribution, I find that for Italy composition effect only affects the upper tail (90-50 gap) earnings distribution.

The fail in attributing to the compositional changes in the labour force the spurious effect in increasing residual wage inequality in Italy and UK, means that there is more to be understood about the interactions among changing characteristics of the labour force, the labour market institutions and the macroeconomic conditions.

For the Italian experience the modest role of the increase in educational attainment on the residual, could be related to the observed decrease in educational premia along with a stable trend in the demand for skills (Naticchioni et al. 2008). The centralised wage structure that characterizes the Italian labour market and that is likely to limit the differential in rewards to skills (both observed and unobserved) may in part explain, along with the low share of graduate workers, why the composition effects fail to exert a spurious effect of the observed increasing within-group inequality.

In both the UK and the U.S. residual wage inequality has been growing in tandem with the share of immigrants in the labour force; immigration is often view as a proximate cause of the rising wage inequality; however the overall impact of

immigration on native inequality depends not only on the effects on between-group differentials but also on the effects on within-group inequality.

Based on this framework chapter three focuses on the effects of immigration on the residual wage inequality in the UK and US between 1994 and 2008, in particular it seeks to assess whether and to what degree immigration contributed, along with technology, institutions and traditional explanations, to widening inequality.

This investigation reveals that the presence of immigrants does not have a causal relationship with the increase in residual inequality. Even when treating natives and immigrants as two separate groups, the trend of residual variance inequality for natives does not significantly change, suggesting that inequality for natives is not due to the unobservable skills of immigrants acting in the same labour market.

When the share of immigrants in the labour force is held fixed at its value in the base year, the role of the residual components ranges from 0.002 to 0.009 for natives and up to 0.008 for all workers. Although the composition effect still explains most of the growth in the residual for natives and the whole labour force, the results also suggest that the presence of immigrants plays a role in explaining the growth in the residual variance observed, albeit a very small share. On average, the non-causal effect of immigration on residual variance ranges between 0.1% and 0.5% of the observed change between 1994 and 2008.

The increasing number of immigrants in the labour market has been a major policy and public concern in both the UK and the U.S. The current economic downturn has added further momentum to what in many countries already represents an important and controversial public policy issue about the impact of rising numbers of immigrants on the labour market outcomes for natives, particularly the impact of immigration on the wages of native workers.

In terms of public policy the findings of chapter three confirming almost no effect of the increase of immigration on wage inequality in the UK and the U.S., imply that the increasing share of immigration in the UK and U.S. should not be a concern in the public policy agenda.

Recently, (January 2010), the National Equality Panel Report outlined that Britain is an unequal country, more than many other industrial countries and more than a generation ago. There are several differences in labour market outcomes and inequality across the English regions with inequality in many dimensions being wider in London than in any other region.

Motivated by the fact that expenditure on housing is the largest component of total household expenditure and varies considerably through regions in the UK, and that graduate workers are more concentrated in more expensive British regions, in chapter 4 I reassessed 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.

This investigation does reveal a number of new points: 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 and 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.

The failure of the National RPI to appropriately reflect the real cost-of-living of different UK regions recalls how 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). The results of chapter 4 also suggest the need for regional specific studies and related policy to address the existing regional differences in

the labour market and standards of living in the UK. For example, London 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).

The increasing gap between the wages of skilled versus unskilled men suggests that the demand for skills has increased, either because of changes in technology or changes in the structure of product markets, and that the supply of skilled labour has not changed fast enough to compensate (Gosling et al., 1996). This may be because of the failure in the education and training system which prevents unskilled workers gaining the skills demanded by the labour market. Therefore, to reduce the skills shortages and improve opportunities for workers at the bottom of the distribution, it is likely that policies that encourage children to stay on at school and more people to go into adult and further education are needed.

A number of analysts have pointed to the way in which large inequalities are associated with societies having lower levels of happiness or well-being in other aspects, and to the social problems and economic costs resulting from these.

Among the general public as well as scholarly researchers, it has spurred a vigorous debate about how to regard widening inequality and what, if anything, to do about it. Although there has been much debate over whether the proper object for public-policy concern is inequality per se or simply the low incomes of those at the bottom of the distribution, either view leads to the conclusion that it would be good to raise the earnings power of the lowest-income workers. Although there remains widespread disagreement about the relative importance of different explanations for the recent increase in inequality and the obstinate persistence of poverty, the role of institutions and labour market skills and workers education figure high on nearly everyone's list. It can be in the interest of the wealthy, as well as the poor, to raise the incomes of the poor. For example, minimum wage legislation has been framed with the intent to compress wage inequality. But the effectiveness of the minimum wage as an income redistribution tool is often criticised, since by raising the cost of labour it can have negative effects on output and employment.

Particular policies can be used to alter both the skill endowment of the workforce and the shape of the wage distribution. There is, for example, some controversy about the role that migrants should play in reducing domestic labour shortages in particular sectors and occupations; this has been investigated by Ruhs and Anderson (2010) who questioned how to link the admission of migrant workers to the "needs" of the domestic labour market.

Moreover, providing potentially low-income workers with more of the kind of "human capital" that matters for labour market outcomes is an avenue for public policy. Growth in the quality of the U.S. workforce has been a major source of productivity growth and economic mobility over the past century.

When considering whether the degree of inequalities is "justified" or not, differentials in wage that reflect differences say in work experience, creating differences by age, might be seen as logical. As claimed by Krueger (2005) "[…] *inequality has both positive and negative effects. On the positive side, differential rewards provide incentives for individuals to work hard, invest, innovate. On the negative side, differences in rewards that are unrelated to productivity are corrosive to civil society and cause resources to be misallocated."*

As outlined by Krueger (2005), societies must strike a balance between the beneficial incentive effects of inequality and the harmful welfare-decreasing effects of inequality. The optimal balance will differ across societies and time, but too much inequality can repress innovation. Expanding education and training programs for less-skilled workers could be an effective component of a strategy to restore a better balance.

References

Abdurrahman, A. and G. Borjas (2007) 'Cross-country variation in the impact of international migration: Canada, Mexico, and the United States', Journal of the European Economic Association, 663-708.

Acemoglu D. (2002) 'Technical change, Inequality, and the Labour Market', Journal of Economic Literature 40(1): 7-72

Ashenfelter, O. and R. Layard (1999) *Handbook of labour economics*. Handbooks in economics. Vol. 5. Amsterdam; Oxford ; North-Holland.

Ashworth, J., and S. C. Parker (1997) 'Modelling regional house prices in the UK', Scottish Journal of Political Economy, 44 (3) (08): 225.

Autor D., L.F. Katz, and M. Kearney (2005) 'Rising Wage Inequality: the Role of Composition and Prices', NBER Working Paper No. 11628

Autor, H.,D., and L.F. Katz (1999) Changes in the Wage Structure and Earnings Inequality. *Handbook of Labour Economics*, Volume 3, North Holland.

Autor, H.,D., L.F. Katz, and M. Kearney (2008) 'Trends in U.S. wage inequality: Revising the revisionists', Review of Economics & Statistics 90, (2) (05),300-23.

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.

Becker, G. (1964) '*Human capital: a theoretical and empirical analysis with special reference to education*', New York , NBER.

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).

Blau, F., D., and L., M. Kahn (1996) 'International differences in male wage inequality: Institutions versus market forces', The Journal of Political Economy, 104 (4) (Aug.): 791-837.

Blau, F., D., and L.,M. Kahn (1996) The gender earning gap: some international evidence. In *Differences and Changes in wage Structures*, edited by Richard B. Freeman and Lawrence F. Katz. Chicago: Univ. Chicago Press (for NBER), 1995.

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.

Borjas, G. J. (2006) 'The impact of immigration on the labour market', Paper prepared for the conference of Labour and Capital flows in Europe Following Enlargement, organised by International Monetary Fund, the Joint Vienna Institute, and the National Bank of Poland.

Borjas, G., J., R., B., Freeman, and L.,F. Katz (1996) 'Searching for the effect of immigration on the labor market', The American Economic Review 86, (2, Papers and Proceedings of the Hundredth and Eighth Annual Meeting of the American Economic Association San Francisco, CA, January 5-7, 1996) (May): 246-51.

Borjas, G.,J., J., Grogger, and G. Hanson (2008) 'Imperfect substitution between immigrants and natives: A reappraisal', NBER Working Papers 13887.

Borjas, G. J. (2008) *Labor economics*. 4th ed. New York: McGraw-Hill Higher Education

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.

Boudarbat, B. and T. Lemieux (2008) 'Why are the relative wages of immigrants declining? A distributional approach', Canadian Labour Market and Skills Researcher Network. WP n.65.

Bound J. and G. Johnson (1992) 'Changes in the Structure of Wages in the 1980s: an Evaluation of Alternative Explanation', American Economic Review 82(3): 371-92.

Brunello, G., Fort, M. and G. Weber (2009) Chnges in Complulsory schooling, education and the distribution of wages in Europe. The Economic Journal, Vol. 119, Issue 536, pp. 516-539.

Burkhauser, R.,V., Feng, S. and S., P. Jenkins. (2007) 'Using the P90/P10 index to measure the US Inequality trends with Current Population Survey Data: a view from inside the Census Bureau Vaults', ISER Working Paper 2007-14. Colchester: University of Essex.

Cappellari, L. (2002) 'The Dynamics and Inequality of Italian Men's Earnings: Long-Term changes or Transitory Fluctuations?', The Journal of Human Resources, vol.39, No.2, pp.475-499.

Card D., and J. DiNardo (2002) 'Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles', Journal of Labor Economics 20(4): 733-83.

Card, D., (2005) 'Is the new immigration really so bad?', Economic Journal 115, (4), F300-F323.

Card, D., (2009) 'Immigration and inequality', The American Economic Review. Vol. 99, No. 2.

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.

Card, D., F. Kramarz, ,and T. Lemieux (1999) 'Changes in the Relative structure of wages and employment : A comparison of the United States, Canada , and France', The Canadian Journal of Economics. Vol.32, no.4 , pp. 843-877.

Card, D., Lemieux, T. and C. Riddel, (2003) "Unions and the Wage Structure", in John T. Addison and Claus Schnabel (eds.) *The International Handbook of Trade Unions*, Cheltenham: Edward Elgar, pp. 246-92.

Card, D. and A. Krueger (1995) 'Myth and measurement: the new economics of the minimum wage', Princeton University Press.

Carneiro, P., and J.J. Heckman (2002) 'Human Capital Policy', in Inequality in America What role for human capital policies?' (ed. Friedman B.M.), The MIT press.

Checchetti, S. (2007) Housing in inflation measurement. VOX. http://www.voxeu.org/index.php?q=node/248

Chewick, B., (1978) 'The effect of Americanization on the earnings of foreign-born men', Journal of Political Economy 86,897-922.

Chiquiar, D., and H. Gordon(2005) 'International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States', Journal of Political Economy 113, (2) (04),239-81.

Clark, K., and J. Lindley (2006) 'Immigrant labour market assimilation and arrival effects: Evidence from the UK labour force survey', IZA Discussion Papers 2228.

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. (2000) 'Inequalities in income and inequalities in health', NBER wp.7141.

Deaton, A. and J. Muellbauer (1980a) Economics and Consumer Behavior. Cambridge: Cambridge University Press.

Dell'Aringa, C. and C Lucifora (2001) 'Inside the black-box: Labour market institutions, wage formation and unemployment in Italy', Rivista di Politica Economica, 90(X-XI), 21-70.

Deschenes, O. (2001) 'Unobserved ability, comparative advantage and the rising return to education in the United States: a cohort-Based approach', Princeton University Industrial relations section, Working Paper No. 465.

Devicienti, F. (2003) 'The rise in wage inequality in Italy. Observable factors or unobservables?',WP Laboratorio Riccardo Ravelli.

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.

DiNardo, J., N.,M., Fortin, and T. Lemieux (1996) 'Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach', Econometrica 64, (5) (Sep.),1001-44.

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).

Dustmann, C., J., Ludsteck, and U. SchÖnberg (2009) 'Revisiting the German wage structure', Quarterly Journal of Economics 124, (2) (05): 843-81.

Dustmann, C., F., Fabbri, and I. Preston (2005) 'The impact of immigration on the british labour market', The Economic Journal 115, (November), F324-41.

Dustmann, C., T., Frattini, and I. Preston (2008) 'The effect of immigration along the distribution of wages', Cream Discussion Paper series 0803.

Fenwick, D, and J. O' Donoghue (2004) 'Developing estimates of relative regional consumer price levels', ONS.

Firpo, S., N.M. Fortin , and T. Lemieux (2009) 'Unconditional quantile regressions', Econometrica 77 (3) (05): 953-73.

Gallin, J. (2006) 'The long-run relationship between house prices and income: Evidence from local housing markets', Real Estate Economics 34 (3) (Fall2006): 417-38.

Goldin, C. and L. F. Katz (2007) 'Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing,' Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 38(2007-2), pp: 135-168.

Goos, M. and M. Alan (2003) 'Lousily and lovely Jobs: The Rising Polarization of Work in Britain', The Review of Economics and Statistics, Vol. 89:1, February 2007, pp 118-133.

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.

Gosling, A., S. Machin, and C. Meghir (2000) 'The changing distribution of male wages in the UK', The Review of Economic Studies, Vol. 67, NO. 4: 635-666 .

Gould, E.,D., and O. Moav (2008) 'When is "too much" inequality not enough? The selection of Israeli emigrants', C.E.P.R. Discussion Papers n. 6955.

Green, D. A., and T. Lemieux (2008) 'The impact of unionization on the incidence of and sources of payment for training in Canada', In *The economics of education and training.*, eds. C. Dustmann, B. Fitzenberger and S. Machin, 267-291, Studies in Empirical Economics. Heidelberg: Springer, Physica-Verlag.

Green, F., and Y. Zhu (2008) 'Overqualification, job satisfaction, and increasing returns to graduate education', Oxford Economic Papers, 1-24.

Gregg, P. and S. Machin (1994) 'Is the UK rise in inequality different', in Barrell, R. (ed.), The UK Labour Market. Cambridge. Cambridge University Press.

Hanson, G. H., (2006) 'Illegal Migration from Mexico to the United States', Journal of Economic Literature, vol. 44(4), pages 869-924, December.

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.

Hills, J., Brewer, M., Jenkins, S., Lister, R., Lupton, R., Machin,S., Mills, C., Modood,T., Rees, T., Riddel,S. (2010) 'An anatomy of economic inequality in the UK: report of the national equality panel'. Government Equalities Office.

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.

Juhn, C., K.,M., Murphy, and B. Pierce (1993) 'Wage inequality and the rise in returns to skill', The Journal of Political Economy 101, (3) (Jun.) 410-42.

Katz, L.,F., and K.M. Murphy (1992) 'Changes in relative wages, 1963-1987: Supply and demand factors', The Quarterly Journal of Economics 107, (1) (Feb.),35-78.

Konus A.A. (1924) 'The problem of the true index of the cost of living', Reprinted in Econometrica,vol.7 page 10-29.

Krueger, A. (2002) 'Inequality, too much of a good thing', in Inequality in America What role for human capital policies? (ed. Friedman B.M.), The MIT press.

Krugman, P. (1994). 'Past and prospective causes of high unemployment', in Reducing unemployment: Current Issues and Policy Options. Proceedings of a Symposium in Jackson Hole, Wyoming (Kansas City, Mo: Federal Reserve Bank of Kansas City).

Kuznets, S. (1955) 'Economic growth and Income inequality', American Economic Review 45:1-28.

Lemieux, T. (2006a) 'Post-secondary Education and Increasing Wage Inequality', American Economic Review Papers and Proceedings 96(2): 195-199.

Lemieux, T. (2002) 'Decomposing Changes in Wage Distributions: A Unified Approach' , Canadian Journal of Economics, 35(4), November 646-688.

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.

Lemieux, T. (2008) 'The changing Nature of Wage Inequality', Journal of Population Economics, Springer vol. 21(1), 21-24.

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.

Levy, F. and R.J. Murnane (1992). 'U.S. earnings levels and earnings inequality: a review of recent trend s and proposed explanations, Journal of Economic Literature 30: 1333-81.

Lilla M. and S. Staffolani (2007)' The Evolution of Wage Inequality in Italy', Economic Letter (2008).

Lilla M., (2005) 'Disuguaglianze salariali in Italia: nuove evidenze dai microdati SHIW', Politica Economica, no.1, pp71-102.

Lucifora, C. (2001) 'Wage inequality and low pay: the role of labour market institutions', in Gregory, M., W., Salverda and S. Bazen (Eds.) Labour market inequalities, ed. Oxford University Press.

Machado J.A.F., and J. Mata (2001) 'Earnings function in Portugal 1982-1994: evidence from quantile regressions', Empirical Economics 26: 115-134.

Machin, S. (1996) 'Wage inequality in the UK', Oxford Review of Economic Policy, 7, 47-62.

Machin, S. (1999) '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. (2008) 'An appraisal of economic research on Changes in wage inequality', Labour, vol.22 p: 7-26.

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. 2010. Changing wage structures: trends and explanations. Working paper.

Machin, S., and J. Von Reenen (2008) 'Changes in wage inequality', New Palgrave Dictionary of Economics.

Manacorda M. (2004) 'Can the Scala Mobile Explain the Fall and Rise of Earnings in Italy? A Semiparametric Analysis, 1977-1993', Journal of Labor Economics , 22(3).

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.

Martins, P.S., and P.T. Pereira (2004) 'Does education reduce wage inequality? Quantile regression evidence from 16 countries', Labour economics, 11, p: 355-371.

Melly, B. (2005) 'Decomposition of differences in distribution using quantile regression'. Labour Economics 12, (4), 577-90.

Mincer, J. (1958) 'Investment in human capital and personal income distribution' Journal of Political Economy, 66(4), 281-302.

Mincer, J. (1974) *Schooling, Experience and Earnings* New York: National Bureau of Economic Research.

Mincer, J., (1997) 'The production of human capital and the life cycle of earnings: Variations on a theme', Journal of Labor Economics 15, (1, Part 2: Essays in Honor of Yoram Ben-Porath) (Jan.),S26-47.

Moretti, E. (2004) 'Workers 'education, spillovers, and productivity: evidence from Plant-level production functions. The American Economic Review, Vol. 94, No. 3, pp. 656-690.

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.

Murphy, K., and F. Welch (1992) 'The structure of wages', Quarterly Journal of Economics, 107, 285-326.

Naticchioni P., A. Ricci, and E. Rustichelli (2008) 'Wage Structure, Inequality and Skill-biased change: is Italy an Outlier?', Labour Vol.22, Issue s1.

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.

Ottaviano, G., and G. Peri (2006) 'Rethinking the effects of immigration on wages.' NBER Working Paper No. 12497.

Peri, G., and C. Sparber (2009) 'Task specialization, comparative advantages, and the effects of immigration on wages', NBER Working Paper no. 13389.

Rienzo, C. and Carlos Vargas-Silva (2010) Briefing paper for the Migration Observatory, University of Oxford.

Ruhs, M. and B. Anderson (2010) "Who Needs Migrant Workers? Labour Shortages, Immigration and Public Policy", Oxford University Press.

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.

Schmitt, J., (2003) 'Creating a consistent hourly wage series from the Current Population Survey's Outgoing Rotation Group, 1979-2002'.

Schmitt, J., and J. Wadsworth (2007) 'Changes in the Relative Economic Performance of Immigrants to Great Britain and the United States, 1980–2000', *British Journal of Industrial Relations*, vol. 45, no. 4 (December), 659-686.

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.

Welsh, F. (1979) 'Effects of cohort size on earnings: the baby boom babies' financial bust', The Journal of Political Economy, vol.87, issue 5, part 2.

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