Mobility, Education, and Skills: A Labour Market Perspective on the UK's Industrial Strategy

Man Yin Leung

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy in Economics

Swansea University

2024

Abstract

This thesis utilises a regional framework to explore the assumptions of the Industrial Strategy regarding skills and their impact on the UK economy. Through three empirical chapters, it examines underlying issues and proposes potential policy developments.

The first chapter investigates how graduate degree choices and mobility decisions affect individuals' earnings and the UK economy, particularly addressing brain drain issues in certain regions. Using the 2004 to 2013 HESA data and the RIF regression technique, the analysis assesses migration patterns' impact on graduates' wage premiums, considering factors like horizontal mismatch and regional cost of living. The results show that except for Move Returner, all the migration types on average, earn a positive wage premium.

The second chapter focuses on understanding the combination of skills needed for job-specific requirements and examines how job mismatches affect wages. The data is extrapolated from the Labour Force Survey for the period between 2012 to 2020. Econometric models analyse how skill and job mismatches impact UK graduates' wage premiums, exploring potential disparities for ethnic minorities. The empirical findings shows that vertical mismatch causes a greater negative wage effect compared to the horizontal job mismatch.

The final chapter examines the development and importance of generic skills for the UK population aged 20 to 65, considering concerns about skill underinvestment. Using data from 2012 to 2017 and Unconditional Quantile Regression, the analysis estimates return based on the importance of generic skills, including the influence of Computer Complexity and its implications for wage returns in the context of technological advancements. Additionally, the chapter explores ethnic minority perspectives on wage discrepancies. The findings shows that Computer Complexity (Physical skills) provides the greatest positive (negative) wage premium effect.

Declarations and Statements

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed: Man Yin Leung (candidate)

Date: 31/03/2024

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s). Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed: Man Yin Leung (candidate)

Date: 31/03/2024

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed: Man Yin Leung (candidate)

Date: 31/03/2024

Table of Contents

Acknowledgements

I wish to express my deepest appreciation to my supervisors, Professor Nigel O'Leary and Professor David Blackaby, for their invaluable guidance, expertise, and unwavering support throughout this journey. Their mentorship has not only influenced my research but has also contributed significantly to my development as a scholar. I am especially grateful for Nigel for guiding me through tough times, both academically and my life difficulties. He has always been super supportive, constructive and communicates in a very friendly way. I could not have achieved this without the immense support.

I am immensely grateful to the Economic and Social Research Council for funding this research endeavour. Additionally, I would like to acknowledge the Economic department within the School of Social Sciences for providing the necessary resources and fostering an environment conducive to research and learning.

Heartfelt thanks go to my family and my partner for their unwavering love, understanding, and encouragement. Their belief in me has been a constant source of strength and motivation.

I am deeply appreciative of the support received from my teammates at Swansea University Badminton team, especially Steve Kuriakose, who have been pillars of support throughout my PhD journey. Also, I am grateful for other badminton members such as Will Freer and Yu Gene.

To my friends who have provided unwavering support, I extend special thanks to Monil Khera, Saskia Davies and Charles Chu. Your shared experiences and encouragement have been instrumental in reaching this milestone. They have helped me mentally through tough and good times. Also, Monil and Saskia have helped with my ideas during my PhD which I am grateful for.

List of Tables

List of Figures

Chapter 1

Introduction

The development and retention of high-skilled labour have become crucial elements in shaping modern economies, particularly within the context of the UK's Industrial Strategy. Launched in 2017, the strategy was created in response to pressing national challenges, such as sluggish productivity growth, regional economic disparities, and the growing demand for skills in an increasingly digital and globalized marketplace. These factors, combined with disruptive events like Brexit and the COVID-19 pandemic, have exacerbated the country's skill shortages, highlighting the urgent need for coordinated efforts to bridge the gap between available talent and economic needs.

One of the central challenges facing the UK is its so-called 'productivity puzzle,' where productivity growth has stagnated since 2010, despite efforts to stimulate economic expansion. Skill shortages, particularly in STEM (Science, Technology, Engineering, and Mathematics) fields, continue to hinder national growth potential. Furthermore, regional disparities have seen certain areas, especially those outside London and the South East, struggling to retain graduates, further weakening regional economic performance.

Research on the role of skills in supporting national productivity and reducing regional inequalities is still in its early stages, and existing studies often focus on narrow aspects of the issue, such as specific industries or geographic areas. However, there is a clear need to build a more comprehensive understanding of how skill mismatches, migration patterns, and regional policies interact to affect labour market outcomes across the UK. Specifically, it is critical to explore how vertical (over- or under-education) and horizontal (misalignment between education and job roles) mismatches affect wages and productivity, particularly among ethnic minorities and regions with pronounced skill deficits.

This thesis seeks to fill these gaps by analysing the UK's skills landscape through a regional lens. By leveraging large datasets and advanced econometric techniques, this research aims to provide actionable insights into how the UK can more effectively develop, retain, and utilize high-skilled labour to achieve more equitable economic growth. Given the rapid changes in the global economy and the rising importance of technological skills, the findings of this study will be pivotal in shaping future industrial strategies and regional development policies, making it both timely and essential for future research.

1.1. Background of the Industrial Strategy

The UK's Industrial Strategy, introduced in 2017, was developed as a response to a range of economic challenges that include stagnant productivity growth, pronounced regional economic disparities, and the need to adapt to technological advancements in a rapidly evolving global economy. This comprehensive policy framework is designed to address structural weaknesses in the UK economy while leveraging the country's strengths to create a more prosperous and equitable society. The strategy consists of multiple pillars, each aimed at fostering productivity and growth through targeted investments and policy interventions. These include enhancing skills development and education, increasing research and development (R&D) and capital investment, strengthening infrastructure, and creating a business environment conducive to innovation and investment. In this context, the term "industrial strategy" is used broadly to encapsulate these initiatives, each contributing to a cohesive policy approach aimed at sustainable economic growth. While this thesis focuses specifically on the skill dimension, particularly as it relates to wages and labour market dynamics, these elements are part of a broader strategy aimed at supporting national productivity, labour market flexibility, and economic resilience.

The primary rationale for the UK's Industrial Strategy stems from the country's persistent productivity puzzle, where productivity growth has stagnated despite various policy interventions. Productivity, or the measure of output per unit of labour, is a key determinant of living standards and economic competitiveness. The strategy's aim is to address this stagnation by investing in the skills, technologies, and infrastructure needed to support a competitive, high-growth economy. Central to this is the role of skills development, as a skilled workforce is foundational to improving labour productivity and enabling effective participation in high-value industries. In recent years, skill shortages, especially in STEM (Science, Technology, Engineering, and Mathematics) fields, have highlighted the urgent need to align the educational system with labour market demands. However, the industrial strategy also recognizes that skill development alone is insufficient; it must be complemented by investments in R&D, infrastructure, and a supportive business environment that attracts talent, fosters innovation, and encourages economic participation across all regions.

The strategy's focus on R&D and capital investment reflects the growing recognition that economic competitiveness relies on continuous innovation. By promoting R&D activities, the strategy aims to create an environment where new ideas can flourish, supporting industries that drive economic growth and technological progress. Such investments are particularly important for sustaining high-value sectors, including manufacturing, advanced technology, and life sciences. The UK government has made a commitment to increase R&D spending as a percentage of GDP, aligning with international benchmarks to ensure that the UK remains competitive in the global innovation landscape. Furthermore, capital investment in technology and automation plays a key role in driving productivity gains, particularly in sectors that have been slow to adopt advanced manufacturing techniques. Capital investment also enhances productivity by reducing costs, improving efficiency, and increasing the quality of goods and services produced.

Infrastructure development is another critical component of the industrial strategy. Infrastructure, including transportation networks, digital connectivity, and energy supply, forms the backbone of economic activity. By enhancing infrastructure, the strategy seeks to create an environment where businesses can operate efficiently, and regions can attract investment. For instance, improved transportation networks reduce travel times and increase accessibility to jobs, thereby supporting labour mobility and helping to address regional disparities. Digital infrastructure is equally important, as it enables businesses to leverage technology, access new markets, and improve operational efficiency. The industrial strategy prioritizes both physical and digital infrastructure to ensure that all regions can participate in the economy and benefit from the growth opportunities created by a connected, technologically advanced society.

The final pillar of the industrial strategy is the creation of a favourable business environment. This involves designing policies and regulations that encourage entrepreneurship, support small and medium-sized enterprises (SMEs), and attract foreign investment. A supportive business environment is essential for fostering innovation and enabling companies to thrive. Policies that reduce regulatory barriers, provide tax incentives, and improve access to finance can create a dynamic ecosystem where businesses can grow, create jobs, and contribute to economic resilience. Furthermore, by focusing on sectoral growth areas such as clean energy, artificial intelligence, and advanced manufacturing, the strategy aims to position the UK as a global leader in emerging industries, reinforcing its competitive advantage in the international marketplace.

Within this multifaceted industrial strategy, the focus of this thesis on skills development and wages contributes to understanding a critical element of productivity improvement. Skills development is essential for creating a flexible, adaptable workforce capable of meeting the demands of a knowledge-based economy. A skilled workforce not only supports the growth of high-value sectors but also enables workers to command higher wages, thereby improving living standards and reducing income inequality. This research, which examines labour market outcomes in relation to graduate mobility, skills mismatches, and wage disparities, offers insights into how skill development aligns with broader productivity and economic growth objectives. By focusing on these labour market aspects, this thesis contributes to a key component of the industrial strategy, specifically how skills contribute to labour productivity and regional economic balance.

However, it is essential to recognize that skills development interacts with other components of the industrial strategy. For example, R&D and capital investment require a highly skilled workforce to translate innovations into practical applications that benefit the economy. Similarly, infrastructure improvements enhance the effectiveness of skills by enabling greater regional mobility, making it easier for skilled workers to relocate to areas with high demand for their expertise. Additionally, a favourable business environment can create demand for skilled workers by fostering sectors that rely on specialized knowledge, such as technology, engineering, and financial services. Thus, while this thesis centres on skills and wages, these aspects are interdependent with other elements of the industrial strategy, each reinforcing the others in pursuit of national productivity and growth.

In conclusion, this thesis aligns with the overarching goals of the UK’s Industrial Strategy by providing a focused analysis of labour market issues related to skills and wages. While skills development is only one dimension of the strategy, it is foundational to addressing productivity challenges and ensuring that economic growth is inclusive and sustainable. By examining the dynamics of graduate mobility, skills mismatches, and wage outcomes, this research offers valuable insights for policymakers seeking to strengthen the labour market component of the industrial strategy. Furthermore, this work highlights the importance of integrating skills development with broader investments in R&D, infrastructure, and business environment improvements to achieve a cohesive, comprehensive approach to economic growth. This thesis thus provides both empirical evidence and theoretical context that can inform future policy developments within the framework of the UK's Industrial Strategy, contributing to a more resilient, productive, and equitable economy.

1.2. Research rationale

The rationale behind this thesis lies in the critical intersection between the UK's industrial strategy and labour market dynamics, particularly through the lens of skills, wages, and migration. With the 2017 UK Industrial Strategy as a backdrop, this research addresses two interrelated challenges: the country's productivity stagnation and the underutilization of graduate skills. This thesis contributes by focusing on labour market outcomes associated with real wages, which are typically associated with labour productivity rather than being direct measures. Through an empirical exploration of wage disparities, migration patterns, and skill mismatches, the research provides insights into the role of human capital as a driver of productivity and economic equity.

The Industrial Strategy underscores the importance of a skilled workforce for improving national productivity and fostering inclusive economic growth across regions. While the policy framework identifies STEM (Science, Technology, Engineering, and Mathematics) skill shortages as crucial, recent data indicate that broader skills gaps—such as those related to interpersonal, technical, and digital skills—pose significant obstacles to realizing productivity gains. This thesis, therefore, adopts a labour market perspective on the Industrial Strategy, examining how graduate mobility and job mismatches influence real wages, which indirectly reflect productivity changes.

Real wages serve as a useful proxy for productivity since they are expected to rise alongside productivity in a competitive market. However, several factors, including geographic mobility, job mismatches, and regional economic disparities, create variations in wage outcomes that may distort this relationship. By analysing real wages across different regions, this thesis provides a clearer understanding of how educational and skill mismatches hinder productivity, particularly for graduates who migrate or experience occupational mismatches. In doing so, this work reveals critical insights about the broader labour market undercurrents that affect productivity indirectly through wages.

This study advances labour market intelligence through robust quantitative techniques, including quantile regression and decomposition analysis. The data sources, such as the Labour Force Survey (LFS), Skills and Employment Survey (SES), and Higher Education Statistics Agency (HESA) data, provide the empirical foundation for understanding how skills and regional mobility influence wage outcomes. By examining the wage effects of job and skill mismatches for graduates and analysing how these outcomes vary across ethnic and regional groups, this thesis offers nuanced perspectives on factors that constrain wage growth, productivity, and economic inclusivity. In the context of the Industrial Strategy, this research's contribution is crucial for addressing the labour market inefficiencies that impede productivity gains and equitable wage growth.

1.3. Human capital, skills, and productivity in the context of Industrial Strategy

The concept of human capital—skills, knowledge, and competencies that individuals bring to the workforce—has long been recognized as a fundamental driver of productivity and economic growth. The UK's Industrial Strategy places significant emphasis on developing human capital, particularly by addressing skill shortages and promoting educational initiatives that align with labour market needs. At the core of human capital theory is the idea that education and skills acquisition enhance individual productivity, which, in turn, positively influences wages and economic growth. This chapter provides a discussion of the importance of human capital and skills, their link to productivity via real wages, and their strategic role within the Industrial Strategy.

Human capital is pivotal to achieving high labour productivity, as it enables workers to perform complex tasks, adapt to technological advancements, and drive innovation. Real wages, which typically rise alongside productivity, reflect the value of these skills in the labour market. The link between human capital and wages underscores the importance of ensuring that workers are equipped with relevant skills and that these skills are effectively utilized. However, skill mismatches—where individuals are either overqualified or lack the specific skills required for their roles—can disrupt this relationship, leading to suboptimal wage outcomes and, ultimately, productivity losses.

The Industrial Strategy addresses these challenges by prioritizing skill development across diverse sectors, aiming to build a workforce capable of driving economic resilience and innovation. Within this framework, skills are not viewed as a static resource but as a dynamic component of productivity growth that must evolve alongside industry demands. For instance, while technical and STEM skills are crucial for high-value industries, the strategy also recognizes the importance of generic skills—such as communication and problem-solving—which are transferable and adaptable across sectors. The strategy's skill development agenda is thus inherently tied to the goal of enhancing productivity and reducing regional economic disparities by equipping individuals with skills that command higher wages in the labour market.

The relationship between skills, wages, and productivity becomes particularly significant when considering the role of graduate employment and mobility. Graduates bring specialized knowledge and skills that contribute to higher productivity levels, yet their wage premiums are contingent on their alignment with job requirements. By focusing on graduate mobility and skills mismatches, this research highlights how the Industrial Strategy's objectives of fostering a skilled workforce and promoting regional prosperity can be undermined by persistent mismatches in the labour market.

Through this analysis, the thesis contributes to understanding how human capital can be leveraged within the Industrial Strategy to address productivity challenges. It highlights the importance of aligning educational outputs with labour market needs and demonstrates the economic implications of skill mismatches, particularly in terms of wage outcomes. Ultimately, this work underscores the need for integrated policy efforts that address both the supply and demand for skills to maximize productivity gains and ensure that the UK's Industrial Strategy achieves its goal of a more inclusive, resilient economy.

1.4. Background and research design of each empirical chapter

This thesis utilises regional framework to investigate the assumptions of the Industrial Strategy in terms of skills and enhance knowledge of potential changes in the demand for skills over time. Three chapters are presented to understand the underlying issues and condition in the UK economy and provide potential developments to aid future policies.

1.4.1. Impact of migration behaviour on wage premium

In the first empirical chapter, the research investigates how graduate degree choice and mobility decision impacts an individual's earnings and by extension, the UK economy. Since there is problems of brain drain in certain regions of the UK, emphasis will focus on explaining the skill issues of highly educated young individuals. Also, the patterns of migration for young graduates are examined, with two migration stages[1]. The ability for cities and regions to retain young graduates is vital for regional economic performance and productivity.

To analyse the migration patterns and how it impacts graduates wage premium, the HESA[2] data will be used for estimations by adopting the RIF regression technique by Firpo et al. (2018). There are three main factors that determine whether graduates choose to migrate which will be included in the estimation as observables. Firstly, the issue of horizontal mismatch, where the field of education does not align with the occupation which affects graduates' mobility and lowering productivity in some regions. Secondly, individual characteristics such as age, household circumstances, and racial background influence the choice of region for HE and subsequent workplace. Lastly, the regions cost of living such as housing costs and availability significantly shape migration patterns.

[1] First migration stage records the movement from the individual's region of domicile to the region of HE studies i.e. university. The second migration stage records the movement from the individual's region of HE studies to the region of workplace.
[2] Students who graduated from HE during the academic years between 2002 – 2013.

1.4.2. Impact of job mismatches on wages

In the second empirical chapter, the focus is on understanding the combination of skills needed to support job-specific requirements for adapting within or transitioning between jobs. Interest in educational attainment levels has surged over the past decades, particularly in recent years as the demand for various skill sets has evolved rapidly. Mason et al. (2018) argues that the investments in graduates could alleviate the issue of skill shortages. In particular, the UK government have concentrated on the growth of STEM skills.

There are two primary mismatches when one enters the labour market. The first circumstance is vertical mismatch which occurs when graduates are either overeducated or undereducated relative to the occupation they are employed. The second is horizontal mismatch which occurs when the graduate's field of study or skills acquired through education does not align with the requirements of the occupation. To address this, econometric models will be utilised to analyse how UK graduates' skill and job mismatches effects their wage premium. The Office for National Statistics (ONS) and Labour Force Survey between the 2012 to 2020 period is used for this empirical chapter. Further research will investigate if the skill mismatch issues are exacerbated for the ethnic minority compared to the White British group.

1.4.3. Investigation of the returns to generic skills

The final empirical chapter focuses on the development of the generic skills for the UK population between the ages of 20 and 65. The Industrial Strategy revealed that the underinvestment in skills is a primary factor for the recent poor productivity in the UK. It is imperative to adjust and adapt the supply of the labour force skills to the constant advancement of the skills demanded by the labour market. Internationally, significant investment has been made in computing and technological progress, areas where the UK must strive to be competitive to drive economic growth and innovation. However, the International Labour Office (2010) argues that for sustainable and balanced growth, it is imperative to implement policies to equip a skilled workforce with a range of essential skills.

To examine this topic, empirical analysis investigates the importance of possessing generic skills and the types of generic skills in the UK labour market. The sample year of this study is between 2012 and 2017. The Unconditional Quantile Regression is utilised to estimate the returns depending on the importance of the generic skill of interest. There is an inclusion of the Computer Complexity variable to further examine the influence on the wage returns, especially since technological advancements have been rapidly growing in recent years. Furthermore, to provide a deeper understanding of the wage discrepancy, the ethnic minority[3] aspect has been attempted is explored.

1.5. Contribution

This thesis contributes to the literature by examining key labour market dynamics within the framework of the UK's Industrial Strategy, particularly through the lens of real wages and skills. While the Industrial Strategy emphasizes several pillars, including infrastructure, R&D, and business environment, this research narrows its focus to labour market issues, specifically the role of skills, wage determination, and migration. By focusing on real wages, which are generally positively associated with labour productivity rather than a direct measure of it, this research offers an empirical investigation into the relationships between skills, wage outcomes, and productivity within the UK's unique economic context.

Real wages are of particular interest in this thesis, as they serve as a proxy for productivity improvements in the labour market. Real wages tend to rise when labour productivity increases, reflecting a higher valuation of the workforce's contributions within the economy. However, the association between real wages and productivity is complex and influenced by multiple factors, including skill mismatches, regional disparities, and migration patterns. This thesis seeks to unpack these layers, exploring how wage outcomes are influenced by the alignment between workers' skills and job requirements, graduate migration patterns, and variations in economic opportunity across UK regions.

The first empirical chapter focuses on the wage effects associated with graduate mobility. Through detailed analysis using a pooled longitudinal dataset and advanced econometric

[3] White British group and the Non-White group counterpart.

techniques, such as the Oaxaca-Blinder decomposition and quantile regression, this research examines whether graduates who migrate for education but later return to their regions face wage penalties compared to those who remain in metropolitan areas. By exploring these patterns, the research highlights the critical role that regional labour markets and mobility constraints play in shaping wage outcomes, adding to our understanding of how labour market dynamics intersect with regional economic performance within the UK's industrial landscape.

The second empirical chapter builds on this by examining the role of skill mismatches in wage determination. This chapter explores both vertical (over- or under-education) and horizontal (misalignment between field of study and occupation) mismatches. The findings demonstrate that such mismatches can significantly affect wage outcomes, which indirectly reflects productivity levels by highlighting inefficiencies in skill utilization. By using extensive survey data from the Labour Force Survey, this analysis provides insights into the factors driving wage disparities, focusing particularly on the additional challenges faced by ethnic minorities within the labour market. This research offers important implications for policy interventions that seek to reduce skill mismatches and improve labour market inclusivity, which in turn can help alleviate productivity constraints in the economy.

The final empirical chapter shifts focus to the returns on generic skills, with a particular emphasis on digital and computer skills, which have become increasingly valuable in the digital economy. By employing Unconditional Quantile Regression (UQR) analysis, this research investigates how these generic skills impact wage premiums across different demographic groups. This chapter also addresses the varying returns on skills across ethnic and gender groups, shedding light on disparities in the labour market. Findings suggest that digital skills yield substantial wage premiums, underscoring the need for policy initiatives that prioritize digital skill development within the workforce. These insights are critical to understanding how skill composition within the labour market impacts wage distribution and productivity outcomes.

By situating these findings within the framework of the UK’s Industrial Strategy, this thesis makes several unique contributions. First, it deepens our understanding of how real wages, as a reflection of labour productivity, are influenced by migration patterns, skill mismatches, and regional economic conditions. Second, it provides empirical evidence that supports the Industrial Strategy's emphasis on skills as a foundation for economic resilience and productivity growth. Finally, it highlights the need for targeted regional policies that address wage disparities and support the development of a more inclusive labour market, where skills are aligned with the evolving needs of a technology-driven economy.

In sum, this thesis offers valuable insights into the labour market aspects of the UK's Industrial Strategy, particularly the critical role of skills in driving productivity through real wage improvements. Through its focus on the dynamics of graduate mobility, skill mismatches, and returns on generic skills, this research provides actionable recommendations for policymakers seeking to align labour market outcomes with broader productivity and economic growth objectives.

Chapter 2

Examining the impact of Migration Behaviour on Wage Premium: Quantile Regression approach on the 2004 – 2012 UK University Graduate Population

2.1. Introduction

The aim of this study is to investigate the influence of graduates' choice of degree and mobility decisions on the UK economy, as well as to assess the impact of the industrial strategy on this decision-making process. Graduates from universities are regarded as valuable assets due to their supplementary education, making them crucial human capital for local and regional development. Consequently, there is a burgeoning interest in understanding cities and regions' capacity to retain young graduates to enhance economic performance. Traditionally, internal migration in the UK is largely motivated by highly educated youth in pursuit of professional opportunities (Champion, 1999).

The existence of universities exemplifies the elevation of human capital levels with a specific region which increases the local supply and demand for skilled labour (Abel and Deitz, 2009). Sequentially, the presence of universities provides a determinant of occupational as well as human capital speciality within one’s region. Therefore, since various regions are generally inclined to specific structures of occupations, only certain regions enjoy the benefits of obtaining highly educated young graduates. This in turn instigates a human capital drain issue whereby cities that are economically worse off, are found to have a lower proportion of young graduate in their workforce; a vicious cycle issue which has a snowball effect (Simmie et al, 2006).

The establishment of universities signifies the advancement of human capital within a particular region, leading to an increase in both the local supply and demand for skilled labour (Abel and Deitz, 2009). Consequently, the presence of universities becomes a determining factor in the specialisation of occupations and human capital within a region. As different regions tend to specialise in specific occupational structures, only certain areas benefit from attracting highly educated young graduates. This phenomenon exacerbates the issue of human capital drain, whereby economically disadvantaged cities have a lower proportion of young graduates in their workforce, perpetuating a vicious cycle (Simmie et al., 2006).

This chapter analyses the persistent wage inequality among young university graduates in the UK from 2004/05 to 2012/13, a period between the years preceding and following the 2007/08 financial crisis. Specifically, the aim is to understand the impact of young individuals' migration decisions before and after pursuing higher education (undergraduate) on their acquisition of human capital. To address this question, the revised Oaxaca-Blinder (OB) decomposition method proposed by Firpo, Fortin, and Lemieux (2018) will be employed, which implements the recentred influence function (RIF) regression. This method allows for the estimation of the effect of covariates on equality measures, such as different percentile intervals. The advantage of this refined approach over the original decomposition lies in its ability to further dissect general distributional measures in a manner akin to the conventional OB technique.

In the literature review, an exploration of factors influencing graduates' decisions, including their choice of degree and migration patterns, and how these factors impact their human capital earnings will be conducted. The first concerns the spatial mobility relating to education-job mismatch, encompassing instances of non-utilization of formal qualifications, mismatch in fields of study, and underutilisation of acquired skills. Such mismatches not only motivate migration but also pose costs to regional and national economies by diminishing productivity and impeding growth.

The second aspect revolves around discerning the push and pull factors guiding migrants' relocation decisions. These factors encompass the conditions prevailing in the region of origin and the characteristics of potential destinations. Additionally, individual circumstances such as prior migration experiences, age, time elapsed since graduation, household situation, and work experiences during university studies are explored as influencers.

Lastly, housing situations, including housing prices and availability, exert significant influence on labour markets. This influence becomes particularly pronounced when comparing salary disparities across regions, which significantly shape migration behaviours and mobility patterns.

To resolve the issue of “brain drain” every local and regional economy attempts to identify the key factors that effectively attract and retain young graduates, particularly given the exponential growth of the graduate labour force in the UK (Gillian Bristow et al, 2011). Government policies play a critical role in shaping the career trajectories and circumstances of graduates in the UK. There has been a noticeable increase in government activity in recent decades, with each region implementing its specific policies. For example, Scottish students are exempt from tuition fees if they study in Scotland, the Welsh government offers graduate intern schemes and provides wage subsidies to businesses hiring new graduates, and England has established special agglomeration hubs in certain regions. All these policies are designed to attract and retain highly skilled graduates.

In recent years, governments have increasingly focused on improving the alignment of graduates with careers that correspond to their degree qualifications. Historically, there has been a persistent issue of mismatch between career paths and graduates or cohorts, leading to employment in non-graduate occupations. For example, research by the Department for Education and Skills revealed that nearly 50% of graduates from the 1998-1999 cohort initially found themselves in careers that did not necessitate a higher education degree, though this proportion significantly decreased to 15% four years post-graduation (Elias and Purcell, 2004). Further exploration of occupational mismatch will be undertaken in Chapter 3.

The subsequent section provides an overview of the data utilised in the study, including the study period, variables incorporated, and the data collection method by HESA. The HESA data will be analysed to elucidate key insights, with additional descriptive statistics to be generated. This will involve delineating the general relationships between focal variables, such as human capital earnings in relation to regional distribution. The primary aim of this model is to capture graduate mobility patterns and their impact on labour market outcomes.

Following this, an econometric model will be introduced, outlining the mechanics of the Oaxaca-Blinder (OB) process used to analyse the relationship between graduate mobility and human capital earnings. The estimation process entails a two-stage procedure, with the initial stage involving the decomposition of selected distributional measures, including the 10th, 25th, 50th, 75th, and 90th percentile intervals.

Finally, the econometric results for both the stages will be analysed with emphasis on the overall and wage structure effect decomposition. This will consider the extent to which graduate migration effects their perceived value in the labour market for both migrant and non-migrant graduates. There will be four different categories of migration types, and each will be compared to its counterpart as well as how it fairs with respect to Non-Mover. The conclusion section contains a summary of the overall result and some policy implications.

The econometric results will be analysed to explain the extent to which graduate migration influences their perceived value in the labour market for both migrant and non-migrant graduates. Four distinct categories of migration types will be examined, comparing each to its counterpart as well as assessing their performance relative to Non-Mover.

In the conclusion section, a summary of the overall findings will be provided along with policy implications. This will offer insights into potential strategies to address issues related to graduate migration and its impact on labour market dynamics.

2.2. Literature review

There are two primary focus to examine existing literatures: the premium on graduate earnings and graduate mobility. This study aims to integrate these disparate concepts by investigating the impact of graduate mobility on the earnings premium compared to their non-graduate counterparts. The first stage is to evaluate the literatures that relates to the earnings premium of graduates and the patterns of graduate mobility.

2.2.1. Graduate wage premium

2.2.1.1. Returns for graduates

Becker et al. (1964) introduced the human capital theory, which suggests a positive association between higher education, skills acquisition, and productivity. This theory reflects in the wage premium received by university graduates compared to those with lower qualifications, such as A-levels, as consistently shown in recent studies. For instance, Britton et al. (2020) found higher returns for both men and women with a degree compared to those without. Belfield et al. (2018) demonstrated that, on average, having a degree increases the wage premium by 6% for males and 26% for females by age 29.

Additionally, O'Leary and Sloane (2005) proposed an alternative approach to understand the wage premium of graduates, focusing on the rate of return on educational investment. Using data from the Labour Force Survey (LFS) between 1994 and 2002, they analysed the lifetime earnings increase for graduates compared to non-graduates.

According to O'Leary and Sloane (2005), male graduates can expect a lifetime earnings increase of £141,539 (net of taxes) compared to men with similar characteristics but with two or more A-levels as their highest education level, indicating a significant financial return for male graduates. Without considering tuition fees, the annual rate of return is 10.1%. Even with an assumed annual tuition fee of £3,000, the rate of return remains substantial at 7.3% per annum.

Boero et al. (2021) investigated the average earnings premium of graduates from the 1990 birth cohort using earnings data from 2015-2016, when cohort members were 25 or 26 years old. The study, based on the Longitudinal Study of Young People in England, focused on full-time employed individuals, resulting in a final sample of 1,733 observations. Among them, 43% held a first-degree qualification. Boero et al. found that the mean (median) earnings for graduates in their sample were £24,977 (£24,000), compared to £21,592 (£19,760) for non-graduates, indicating an overall earnings premium of 16% (21%).

However, concerns were raised about potential bias in the survey data due to misreporting of earnings and qualifications. While there may be an upward bias in the estimates of the degree's return using OLS, Battistin et al. (2014) found similarities between transcript and self-reported qualifications data in the National Child Development Study at age 23. Additionally, Dearden (1999) suggested that reporting errors decrease over time between qualification and survey, enhancing the accuracy of observations.

Moreover, Boero et al. used the Longitudinal Education Outcomes (LEO) dataset to examine gross graduate earnings. LEO, an administrative data source, links UK education, benefit, and tax records. Using LEO data, Boero et al. found that 2010/11 graduates, three to five years after qualifying, had median gross earnings of £22,500 and £25,000, respectively.

2.2.1.2. Degree subject choices

Sloane et al. (2005) discovered that male graduates in Mathematics and Computing fields experience the highest lifetime earnings increase, with an estimated premium of £222,419. This corresponds to an annual rate of return of 17.3% and 12.2% when tuition fees are zero and £3,000 per annum, respectively. However, not all degree subjects yield a net lifetime income increase. For example, male graduates in Arts subjects saw only a £22,458 lifetime increase. Adjusted for an estimated cost exceeding £34,000, this results in a negative rate of return between -2.0% and -3.0% per annum.

Similarly, female graduates follow a comparable pattern. Sloane et al. calculated £157,982 lifetime earnings increase for female graduates, translating to a rate of return of 15.0% and 10.3% per annum, respectively, excluding and including tuition fees. This suggests that females generally benefit more in terms of rate of return from attending university compared to males.

Examining differences between genders, female graduates in Education, Mathematics & Computing, and Medical Related subjects would see lifetime earnings increases of £244,740, £227,939, and £208,021, respectively. Even after factoring in the £3,000 annual tuition fees, the returns remain at 17.4%, 15.9%, and 14.4% per annum, respectively. Despite Arts yielding the lowest rate of return for females, their lifetime earnings would still increase by

£113,185, exceeding the costs of attending university, with a rate of return of 7.0% per annum.

Britton et al. (2020) corroborate earlier findings by indicating that, overall, females experience a higher return to education at age 29 compared to males. However, their research reveals that while males see an increase in returns with age, females may experience a decrease in returns for certain degree subjects as they age. Additionally, Britton et al. (2016) found that graduates in Medicine and Economics achieve the highest earnings return regardless of institutional attended or personal characteristics.

Moreover, Belfield et al. (2018a) noted significant variation in returns across different subject choices for both genders. Subjects such as Economics, Medicine, and Law consistently offer higher returns, while Computing, Education, and Pharmacy fall in the mid-range. Similar to previous studies, certain subjects like Arts yield a negative rate of return compared to individuals with A-level qualifications.

Walker and Zhu (2011) estimated the wage premium of graduates by categorizing 12 subject areas into four broad groups: STEM (Science, Technology, Engineering, Mathematics, and Medicine), LEM (Law, Economics, and Management), OSSAH (other social sciences, arts, and humanities including languages), and COMB (degrees combining multiple subjects). Their findings reveal that the premium for male (female) STEM graduates is 25% (38%)[4], LEM graduates is 33% (42%), OSSAH graduates is 10% (33%), and COMB graduates is 20% (33%).

This suggests that the wage premium for LEM subjects tends to be higher on average than for STEM subjects[5]. Moreover, regardless of the chosen subject group, females consistently receive a substantially higher wage premium compared to males. This observation aligns with findings from previous studies. Interestingly, the wage premium for a male studying an LEM subject is comparable to that of a female studying an OSSAH subject.

[4] The graduates wage premium is compared to individuals with 2+ A-Levels.
[5] The graduates wage premium is compared to the baseline for each gender. Females exhibit a lower wage baseline before the inclusion of graduate qualification. Thus, the wage premium from graduates is compared to the specific genders.

2.2.1.3. Degree class

Walker and Zhu (2011) found that achieving a First Class or Second-Class Honours degree, regardless of the chosen subject, results in a substantial positive return compared to other degree classes. This highlights a strong positive correlation between the level of effort and the resulting return. For example, analysing data from the LFS between 1994 and 2004, they discovered that males with an upper second-class degree earn 8% more, and females earn 6% more, compared to those with a lower second-class degree. Similarly, the wage premium for a first-class degree over an upper second-class degree is 4% for males and 5% for females.

Their analysis also extends to postgraduates, revealing that individuals with undergraduate degrees experience an average hourly wage increase of 20% for males and 31% for females compared to those with 2+ A-levels. This premium further increases by 12% for males and 17% for females with a master’s degree, and by 4% for males and 7% for females with a PhD, compared to those with an undergraduate degree. However, individuals with a PGCE experience a decrease (increase) in wage premium by 6% (7%) for males (females) compared to those with an undergraduate degree. This suggests that as individuals progress in their higher education qualifications, the wage premium compared to the previous qualification increases, albeit at a decreasing rate.

Thus, pursuing a postgraduate degree leads to significantly higher returns, regardless of the undergraduate degree class, making further studies worthwhile. However, the benefits of advanced studies diminish with a doctorate. Moreover, Walker and Zhu's findings align with the industrial strategy, suggesting that the UK has a competitive advantage in LEM (STEM) subjects. Therefore, it's advisable to encourage young students, particularly males, to pursue LEM (STEM) subjects during higher education to develop expertise in these areas.

2.2.1.4. Regional effects

The geographical conditions in various UK regions significantly influence the wage premium of graduates and their decisions regarding mobility during and after higher education (HE) studies. O'Leary and Sloane (2008) examined decision-making during the transition from education to employment in the labour market. They analysed 1992 Office of National Statistics (ONS) data to calculate the private rates of return available to university graduates across different UK regions. The results revealed substantial variations in earnings benefits for both male and female graduates across regions. Predictably, London showed a significantly higher rate of return on nominal earnings compared to other regions, except for the South East. However, when accounting for regional living costs, the differences in returns between regions diminished significantly. The South West stood out as experiencing the lowest real rate of return for graduates of both genders, especially when compared to London, largely due to its high cost of living.

Moreover, the diversity of jobs was found to have minimal impact on the relative regional prospects of graduates. Despite differences in occupations and industries across regions, a similar distribution of employment would not significantly affect the average regional earnings of graduates. Regarding the public sector, O'Leary and Sloane suggest that it plays an important role as a labour market area for graduates outside of London. This is attributed to productivity spillover effects, especially in areas with high concentrations of graduates, and the benefits of agglomeration effects.

2.2.1.5. International studies

Comparing graduate returns across countries offers valuable insights into their development trajectories and policy approaches. Most studies focus on Western Europe and North America, but a study from China adds another dimension and potential contrasts in results.

Trostel et al. (2002) conducted a comprehensive study examining the impact of education on wage returns in 28 countries. Using data from the International Social Survey Programme spanning from 1985 to 1995, they estimated that the global average rate of return to schooling is approximately 5% for males and 6% for females for each additional year of education, based on Ordinary Least Squares (OLS) regression analysis.[6]

Several papers have examined the returns of graduates in Germany. For instance, Kamhofer and Schmitz (2015) found a significant increase in the return to higher education, with a positive return of 6.9% for each additional year of higher education. However, they noted no significant difference in returns for compulsory education. Similarly, Ammermuller (2005) reported that males graduating in Business, Economics, Engineering, and other Science degrees earn approximately 90% more income compared to their less educated counterparts, confirming a substantial wage premium for German graduates.

In Bol and Heisig (2021) found significant differences in returns depending on the subject chosen. Subjects such as Medicine, Law, Economics, and Engineering showed the highest returns, while Education, Social Sciences, and Humanities exhibited lower returns. These findings align with trends observed in the United States, where subjects like Medicine, Engineering, Economics, and Law yield higher returns compared to Humanities and Arts subjects.

However, Chen et al. (2020) argues that in China, there has been a decline in the wage premium for graduates due to education expansion projects aimed at modernizing China's educational system by 2035. This influx of new graduates has flooded the job market, leading to a surplus of labour supply and subsequently a decline in the graduate wage premium. Nonetheless, graduates from top-rated institutions studying subjects like Law, Economics, and Management still enjoy a significant wage premium.

Overall, comparing findings from various countries suggests that pursuing undergraduate degrees generally leads to an increase in wage premium. This effect is particularly pronounced in fields like Medicine, Law, and Economics, as well as other subjects to a lesser extent, such as Mathematics, Engineering, Pharmacology, and Education (Britton et al., 2020). Conversely, subjects like Creative Arts and Literature tend to yield lower wage premiums or even negative effects on wages.

[6] This study does not only focus on graduates. Instead, it is measured with the total amount of years in education.

2.2.2. Theoretical background: Migration and labour market flexibility

Migration plays a critical role as a mechanism for labour market adjustment, where regional or sectoral imbalances in wages can prompt individuals to migrate in pursuit of better economic opportunities. Theoretical models such as the Todaro migration model and the Lewis two-sector economic growth model offer essential insights into the dynamics of migration and its role in shaping labour market outcomes.

The Todaro migration model (Todaro, 1969) provides a foundational framework for understanding rural-to-urban migration, particularly in developing economies. The model postulates that migration decisions are influenced by the *expected income differentials* between rural and urban areas rather than current income disparities. Individuals are willing to migrate to urban areas despite high unemployment rates, provided that the expected wage in urban sectors, factoring in the probability of employment, exceeds the rural wage. This theoretical perspective is crucial for understanding migration as a rational response to labour market conditions, even in the face of uncertainty. The implication for labour market flexibility is that migration functions as an adjustment mechanism, redistributing labour from lower-wage rural sectors to higher-wage urban sectors, thereby promoting equilibrium over time.

Similarly, the Lewis two-sector growth model (Lewis, 1954) emphasizes migration as a response to wage differentials between the traditional agricultural sector and the modern industrial sector. In the Lewis model, labour moves from the overpopulated agricultural sector, where marginal productivity is low or zero, to the industrial sector, where wages are higher due to capital accumulation and technological advancements. The industrial sector can absorb this labour at a fixed wage until it reaches a point of labour surplus exhaustion, leading to wage increases. This framework highlights the critical role of migration in fostering economic growth and structural transformation by reallocating labour from less productive to more productive sectors, thus enhancing overall economic efficiency.

In the context of the empirical analysis conducted in this thesis, the independence of the migration variable assumed in the regressions must be discussed in conjunction with these theoretical models. While the Todaro and Lewis models suggest that migration is inherently linked to wage differentials and employment prospects, the regression analysis in this study assumes that migration behaviour can be treated as an independent variable influencing wage outcomes. This assumption is critical for isolating the effect of migration on wage premiums but must be acknowledged as a simplification of more complex migration dynamics, where feedback loops between migration, wages, and employment opportunities could exist.

Understanding migration as both an outcome of labour market flexibility and a mechanism for labour market adjustment enriches the interpretation of the empirical results. By acknowledging the theoretical underpinnings provided by models such as those by Todaro and Lewis, we can better contextualize the observed migration patterns in the UK labour market, where wage disparities across regions may drive migration decisions, ultimately influencing labour market outcomes.

2.2.2.1. Graduate mobility

This section focuses the determinants of graduate mobility between regions in the UK. Graduate mobility holds significant importance as it plays a pivotal role in shaping labour market dynamics, fostering economic growth, driving innovation, and potentially reducing financial disparities between regions. Consequently, governments and policymakers prioritize initiatives aimed at enhancing graduate mobility by supporting education and training programs, formulating immigration policies, and making strategic investments in infrastructure and amenities to attract and retain highly educated individuals.

There exists a spectrum of considerations regarding graduate mobility. While many previous studies focus on a single mobility pathway, it's essential to recognize two distinct movement scenarios: the transition from domicile to university and from university to employment. This research will concurrently examine both movements. Initially, the emphasis will be on identifying the primary factors and constraints influencing migration decisions. Subsequently, attention will be directed towards exploring regional outcomes and the choice of degree subjects among graduates. Lastly, a few studies that analyse both movements will be reviewed, aligning closely with the primary research objective.

2.2.2.2. Determinants and type of mobility

Previous research has extensively explored the initial stage of graduate mobility, which involves the transition from one's domicile region to the university region. A study by Denzler and Wolter (2011) delves into Switzerland, examining how proximity to a university impact both university selection and subject choice. Notably, they focus on Swiss universities where tuition fees are either free or very low, thus minimizing the influence of university decisions on student choices. Their findings reveal that students from socioeconomically privileged backgrounds face negligible constraints in their choice of study. This is attributed to the increased cost of studying associated with greater distance between domicile and university regions, resulting in a correlation between financial background and migration distance for university studies. Moreover, individuals with higher educational attainment are less affected by the distance between domicile and university regions.

Similarly, Venhorst et al. (2011) investigate graduate mobility in a Western European context, particularly in the Netherlands. They assess the spatial mobility of recent Dutch and university graduates and ascertain that graduate migration in the Netherlands is primarily influenced by the spatial distribution of suitable job opportunities. Graduates tend to migrate to regions with more job prospects, with college graduates also considering regional differences in living costs and university graduates focusing more on cyclical factors such as regional economic growth and unemployment rates.

Transitioning to studies focusing on the second stage of migration, Winters (2012) employs microdata from the 2000 Census to examine wage and employment disparities between young college graduates who remain in their college region for study and those who migrate elsewhere for employment post-graduation. Winters finds that graduates tend to settle in regions offering the highest utility relative to their educational achievements. However, graduates from colleges located in smaller regions may face employment disadvantages if they opt to remain in those areas for work, resulting in lower wages, acceptance of jobs in less-educated occupations, and vertical and horizontal job mismatches.

Similarly, Mellander et al. (2011) utilise US data from the Gallup Organisation surveying approximately 28,000 respondents across all 50 states. They investigate the factors influencing individuals' decisions to stay in their domicile region, including economic and demographic characteristics, community economic conditions, and community satisfaction with quality of life. The study finds that community quality of place characteristics significantly outweighs other factors, with physical settings and opportunities for social interaction playing crucial roles in determining residential decisions. While job opportunities hold importance, the influence is comparatively lower, indicating that amenities and quality of life significantly impact in-migration and out-migration patterns, fostering growth in well-established amenity regions.

2.2.2.3. Regional outcomes

There are various methods of categorising regional outcomes, one being the distinction between rural and urban areas. Rothwell et al. (2002), along with other regional and agricultural economists, are concerned about the depletion of human capital in rural regions, noting a decline in population, particularly among individuals with higher levels of human capital. This phenomenon is observed globally, with younger and more highly educated individuals often migrating to urban regions where wage premiums are typically higher (Glaesar and Mare, 2001; Gould, 2007).

Examining specific cases, Corcoran et al. (2010) studied the spatial movement patterns of Australian university graduates, finding high retention rates in urban cities and only a small percentage relocating to remote regions. Similarly, Ahlin et al. (2014) investigated graduate migration in Sweden and observed a trend of young graduates moving to urban areas for employment due to the wider spectrum of job opportunities and higher earning potentials available in urban regions. Their analysis also indicates that graduates with higher academic achievements and more educated parents are more inclined to work in urban settings.

In the context of Wales, Bristow et al. (2011) explored graduate migration and retention, revealing a low retention rate compared to other UK regions. Wales experiences a net export of both students and graduates, with a significant proportion choosing to study or work elsewhere. This trend is consistent with findings from Italy by Dotti et al. (2013), indicating that the ability to attract and retain young graduates is heavily influenced by labour market conditions in both the domicile and potential migration regions.

Bond et al. (2008) propose three main factors influencing graduate migration behaviour: existing connections with different geographical areas, perceived job opportunities, and desired future living conditions. Moreover, regional economists often focus on areas with clusters of high-technology firms, such as Silicon Valley in the USA, to understand the gravitational pull for innovation (Larsen and Wigand, 1987). Castells and Hall (1994) explore high-growth performance sectors to replicate conducive economic and environmental conditions in less attractive regions. Faggian and McCann (2008) suggest that a region's innovativeness is crucial for attracting university graduates, as the influx of young graduates can further stimulate regional innovation, creating a cycle of benefits.

2.2.2.4. Graduate subject degree

Subject degree choice can indeed influence a graduate's decision to migrate. For instance, Faggian et al. (2007a) observed that graduates in Science or Social Science (STEM) subjects are more inclined to migrate for employment compared to Arts graduates. This trend may stem from the fact that Arts graduates often face fewer specific employment requirements and may have opportunities facilitated by university-sponsored placements, reducing the need for secondary migration.

Moreover, factors such as degree class and university ranking can significantly impact migration decisions for employment. Mosca and Wright (2010) found that male graduates with higher degree classifications are more prone to migrate for employment, especially if they hail from more prestigious universities and have already migrated for their studies.

However, Faggian et al. (2007a) present contrasting findings, indicating that British female graduates exhibit a higher likelihood of migration for employment compared to males. This aligns with similar observations in studies conducted in Italy by Congilio and Prota (2008) and in the Netherlands by Venhorst et al. (2011).

2.2.2.5. Ethnicity

Several studies have investigated the extent of graduate migration across different ethnic groups in the UK. Brophy (2021) conducted research on the 2018/19 cohort and concluded that ethnic minority graduates tend to seek employment in more ethnically diverse regions compared to their White counterparts. This aligns with previous studies highlighting the spatial clustering of ethnic minorities in the UK (Longhi, 2014; Zwysen and Demireva, 2020).

Moreover, Brophy (2024) found that graduates from Black and Mixed backgrounds are more inclined to migrate for employment away from their domicile region compared to White ethnic counterparts. This trend is exacerbated by the brain drain effect induced by Greater London. However, these findings contrast with studies by Faggian et al. (2006), which found that Black ethnic groups, particularly graduates, are less likely to migrate for employment. More recent research, such as Donnelly and Gamsu (2018), indicates similar migration rates between Black and White ethnic graduates. On the other hand, Asian and Other ethnic groups are shown to be less likely to migrate for work than their White counterparts (Faggian et al., 2006).

Finally, Brophy (2024) addressed endogenous interaction bias in migration estimation and found that graduates from Asian, Black, and Mixed backgrounds are generally more inclined to leave their domicile regions for employment locations with higher levels of ethnic diversity. Gamsu et al. (2019) further explain that the migration patterns of ethnic minorities remain consistent across both stages of migration, namely entering higher education and the labour market.

2.3. Data

2.3.1. Higher Education Statistical Agency (HESA)

The analysis is based on longitudinal data collected by the Higher Education Statistical Agency (HESA) on students who graduated from UK higher education institutions (HEIs) during the academic year between 2004/05 and 2012/13. This time frame was selected primarily due to the availability of the dataset when the research was conducted. For the 2012/13 Destinations of Leavers from Higher Education (DLHE) Longitudinal Survey, the students graduated in the 2012/13 academic year, which means they typically completed their studies between summer 2012 and summer 2013. The longitudinal survey then followed up with these graduates around 3.5 years later, approximately in 2016.

While more recent data might offer additional insights, the scope of this analysis is confined to the graduates from this specific period. This ensures that the results and conclusions are based on the most comprehensive and accessible data available during the study. Any future extensions of this analysis could incorporate data from subsequent years, should it become available. However, for the purposes of this research, the 2004/05 to 2012/13 dataset provided a solid foundation for exploring the relevant trends and outcomes.

This dataset period is interesting as it captures various economy status and phases. By examining this dataset, insights can be gained into how student decisions respond to fluctuations in the market. Over the period from 2004/05 to 2012/13, the UK economy experienced an average annual growth in labour productivity of 1.6%. The analysis is based on longitudinal data collected by the Higher Education Statistical Agency (HESA) on students who graduated from UK higher education institutions (HEIs) during the academic year between the 2004/05 and 2012/13, there were instances of negative economic growth during 2007/08 and 2012/13.

The main advantage the HESA data lies in the substantial sample sizes obtained through the data collection process, enabling a comprehensive examination of graduation migration patterns. Previous studies have leveraged these datasets due to their size and completeness, as evidenced by works such as those by Hoare and Corver (2010) and Mosca and Wright (2010). Numerous researchers have explored graduate migration using HESA data, as exemplified by Kidd, O'Leary, and Sloane's 2014 study, which investigated whether students who exhibit greater mobility in terms of institution choice and employment location experience a higher wage premium compared to less mobile students. However, this research is limited to the 2003/04 academic year. Nonetheless, the present study aims to extend this research by examining graduate mobility in relation to earnings over multiple years, assessing whether patterns evolve due to both internal and external factors.

The data is compiled by integrating information from the Destination of Leavers from Higher Education (DLHE) survey conducted by HESA. The analysis is conducted in two stages, with the first stage involving a census from Higher Education Institutions. This stage utilizes an administrative dataset that captures background details of individuals who have enrolled in and completed higher education courses in the UK. The recorded information includes personal characteristics (such as age, gender, and ethnicity), course details (like level of study and subject area), and the location of parental or guardian domicile.

To gather this information, HESA annually collects data on graduates' employment status both 6 months and 3.5 years after graduation. The DLHE survey targets all graduates domiciled in the UK and EU from the preceding year, posing inquiries about employment circumstances, annual salary, employment sector, occupation, and employment location.

It's important to acknowledge that HESA excludes certain groups of graduates from its target population. These exclusions encompass graduates with qualifications from further education level institutions, those who primarily studied abroad, exchange students subject to immigration regulations, students undertaking intercalated courses during the recording period, and deceased students. Additionally, HESA continuously refines its data collection methods, which can impact the consistency of datasets as samples may vary in terms of certain variables.

The second phase involves a follow-up survey derived from the DLHE, conducted six months after students graduate. This follow-up survey, known as the Longitudinal Destinations of Leavers (LDLHE) survey, collects graduates' information for up to 3.5 years after they obtain their degree. Unlike the DLHE, the LDLHE is not considered a census survey; instead, it relies on a randomly selected sample from the students who responded to the initial survey. Like the DLHE, the LDLHE offers insights into individuals' personal characteristics, the higher education institution attended, the degree course studied, and employment circumstances such as occupation and salary. Notably, the annual salary data in the dataset is rounded to the nearest thousand pounds before tax.

As HESA tracks the locations of graduates at specific intervals, it allows for the analysis of graduate mobility. HESA records three distinct locations: The domicile of parents or guardians before entry to higher education; The location of the higher education institution attended; The location of employment 3.5 years after graduation. Although the DLHE survey encompasses individuals who graduated from both undergraduate and postgraduate degrees, the analysis in this research will exclusively focus on undergraduates, aligning with the research's objectives. This approach facilitates a less biased and more transparent comparison for a thorough examination of labour market outcomes. Numerous studies strongly suggest that there is a wage premium associated with obtaining a postgraduate degree compared to an undergraduate one (Walker and Zhu, 2008). Therefore, including postgraduates in the analysis would typically overestimate the wage premium for the category that includes a high proportion of postgraduates.

2.3.2. Descriptive statistics

There are many factors which effects the decision of an individual to migrate/non-migrate and the amount one will earn after their higher education (HE) studies. Firstly, the sample sizes of individuals who studied in higher education in terms of domicile, institution, and employment for each year and each specific region of the UK will be depicted.

Table 2.1 presents a comprehensive overview of the distribution of individuals who pursued higher education (HE) in the UK from 2002 to 2012. The Table consolidates data from all these years into a single summary. There are 12 regions considered in this analysis, with regions outside the UK classified as "Outside UK".

In the first column, "Domicile" indicates the permanent residency of graduates before attending HE, with a total sample size of 256,212. The results align with population density patterns, with a significant concentration of students from London (19.18%), followed by the South East (11.39%) and the North West (11.19%). This pattern mirrors findings from previous studies (O'Leary and Sloane, 2017). Conversely, the North East and East Midlands exhibit the lowest domicile figures at 3.70% and 4.40%, respectively.

The second column of Table 2.1 shows the distribution of where individuals studied, Higher Education Institution (HEI) within UK during the 2002 to 2012 period. The graduates chosen location of HEI during their undergraduate studies, on overall, the pattern is very similar to the domicile figures. The two HEI regions with the highest student counts are yet again London and South East with figures of 16.38% and 13.89% respectively. This indicates that London is a net exporter of students as the domicile (19.18%) is greater than the institution (16.38%).

Table 2.1

Distribution of UK students and working graduates by regions (%)

Region	**Domicile**	**Institution**	**Employment**
North East	3.70	4.80	2.97
North West	11.19	10.79	12.01
Yorkshire and the Humber	7.09	8.79	7.04
East Midlands	4.40	6.79	5.84
West midlands	9.09	8.39	7.14
East of England	8.19	5.89	4.85
London	19.18	16.38	24.38
South East	11.39	13.89	9.71
South West	7.39	6.89	7.14
Wales	5.89	6.10	2.67
Scotland	7.79	8.19	7.63
Northern Ireland	4.69	3.10	3.66
Outside UK	0.00	0.00	3.96
Observations	**256,212**	**256,844**	**222,819**

Note: All figures are weighted using *finalwt* to be nationally representative.

The second column illustrates where individuals studied within the UK, referred to as Higher Education Institutions (HEIs), during the same period. The distribution closely resembles domicile figures, with London and the South East being the top HEI locations, accounting for 16.38% and 13.89%, respectively. London experiences a net loss of students during this period, consistent with Swinney and Williams (2016), attributed to factors like high living costs. Conversely, Scotland and Wales see an increase in student numbers due to their unique policy regimes, such as abolishing tuition fees in Scotland and subsidising fees in Wales.

The third column displays the employment locations of individuals three years after graduating from their undergraduate studies. The overall sample size for "Employment" is 222,819, lower than "Domicile", partly due to unemployment. London emerges as a significant employment hub, attracting graduates despite initial domicile figures. Conversely, regions like Northern Ireland and Wales experience a net loss of graduates, highlighting challenges in retaining high-skilled labour.

Wales, in particular, exhibits a concerning trend of losing a substantial portion of its high-skilled labour force post-graduation. This phenomenon is consistent with previous research by Drinkwater and Blackaby (2004), indicating a need for policies to retain Welsh graduates and create high-quality job opportunities locally.

Table 2.2 presents the distribution of individuals across the four UK countries before, during, and after their higher education studies within the UK from 2002 to 2012. England dominates all categories, with figures of 81.63%, 82.58%, and 81.92% for Domicile, Institution, and Employment, respectively. Conversely, Northern Ireland and Wales show declines, particularly in the Institution or Employment phases.

For Northern Ireland, there is a notable decrease from Domicile (4.71%) to Institution (3.14%), though it slightly recovers in the Employment phase (3.68%). This aligns with previous research by Bristow et al (2011), indicating Northern Ireland's status as a net exporter of students to HEIs but retaining some graduates in the employment stage.

Table 2.2

Distribution of UK students and working graduates by countries (%)

Region	Domicile	Institution	Employment	% Change (Dom-Ins)	% Change (Dom-Emp)
England	81.63	82.58	81.92	1.16	0.36
N. Ireland	4.71	3.14	3.68	-33.33	-21.87
Scotland	7.79	8.18	7.73	5.01	-0.77
Wales	5.88	6.13	2.72	4.25	-53.74
Outside UK	0.01	0.02	4.03	N/A	N/A
Observations	**256,212**	**256,844**	**222,819**		

Note: All average distances moved to university are based upon the region of domicile.

Similarly, Wales retains students during HEI studies but experiences a migration trend post-graduation. Both Northern Ireland and Wales are net exporters of graduates, unable to retain their own students or attract those from other UK regions. This concurs with findings by Bristow et al (2011) and Bond et al (2008), highlighting factors such as existing connections, job opportunities, and desired living locations influencing graduate migration behaviour.

The data in Table 2.2 may be skewed by the "Outside UK" category, affecting employment figures as it captures UK students leaving but not international students entering for employment. England shows positive changes in both migration stages, indicating its ability to attract high-skilled labour due to its high concentration of HEIs and relatively high average wage rate.

Scotland maintains a high retention rate for students during HEI studies due to policies like tuition fee coverage by SAAS. However, it becomes a net exporter of graduates in the employment phase, outperforming Wales and Northern Ireland in both movements. This poses challenges for Wales, as highlighted by Drinkwater and Blackaby (2004), regarding the migration tendency of young, educated individuals.

Notably, the total number of students studying in UK HEIs exceeds the number of UK domicile students, suggesting a net inflow of students into the UK. However, there is a significant net outflow of UK graduates working in the UK, potentially due to challenges in obtaining work visas for non-EU students and difficulty retaining international graduate talent, as noted by the Metcalf (2013).

Table 2.3 displays the average distances travelled by students across the twelve UK regions for two migration stages: from domicile to higher education institution (HEI) and from HEI to employment. This analysis aligns with migration theories such as Sjaastad (1962), which view migration as an investment in human capital.

Table 2.3

Distance travelled by students for both migration stages

Region	**Domicile-Institution**	**Institution-Employment**
North East	76.75	97.21
North West	68.63	91.84
Yorkshire and the Humber	72.38	90.87
East Midlands	73.95	83.33
West midlands	69.92	107.90
East of England	112.34	111.10
London	66.12	70.34
South East	112.15	102.47
South West	124.58	113.90
Wales	75.94	83.76
Scotland	78.82	138.17
Northern Ireland	153.98	73.34
Aggregate UK	**90.46**	**97.02**

On average, students travel 90.46 miles from domicile to HEI and 97.02 miles from HEI to employment. Previous research, like that of Faggian, McCann, and Sheppard (2007), predicts an increase in distance for the second migration stage, influenced by factors such as expanding job search areas and higher long-run real wages.

There is a correlation between the two migration stages, with regions showing lower distances for the first stage also exhibiting lower distances for the second stage. For example, North East, North West, Yorkshire and The Humber, East Midlands, London, and Wales have relatively low migration distances for both stages, suggesting a strong correlation between migration history and propensity to migrate.

Conversely, regions like East of England, South East, and South West show higher distances for both migration stages, indicating a positive correlation between previous and subsequent migration. Anomalies are observed in regions like West Midlands, Scotland, and Northern Ireland, where migration behaviour deviates from general patterns due to factors like lifestyle choices and job prospects.

In Scotland, lower migration distances for the first stage are attributed to the government's tuition fee subsidies, but graduates often migrate elsewhere for employment, leading to higher second-stage migration distances. Similarly, Northern Irish individuals may study outside the region but return for employment due to economic benefits, resulting in lower second-stage migration distances.

Overall, the data highlights the complex interplay of factors influencing student migration decisions, including economic opportunities, government policies, and individual preferences.

Table 2.4 presents the aggregate percentages of different movement types adopted by individuals, categorized based on their decisions to leave their domicile region for university and subsequent employment locations. This approach, adapted from Kidd, O’Leary, and Sloane (2014), focuses on movement activity before and after attending higher education institutions (HEIs) across the 12 UK regions, rather than using a 25KM mobility threshold.

The data reveals that 42.16% of individuals across the UK are classified as non-Mover, meaning they study and work in the same region as their domicile. However, this varies

Table 2.4

Graduate mobility patterns by government office region of domicile (%)

Region	Non-Mover	Move-Returner	Stay-Mover	Move-Stayer	Non-Returning Double-Mover	Total
North East	44.41	15.76	15.49	7.03	17.32	100
North West	43.15	17.50	13.05	8.23	18.08	100
Yorkshire and the Humber	43.51	21.39	9.29	7.19	18.62	100
East Midlands	33.99	23.96	9.83	8.86	23.37	100
West midlands	32.13	19.66	10.74	9.19	28.27	100
East of England	13.99	18.86	10.56	12.91	43.68	100
London	44.38	38.73	5.29	4.20	7.41	100
South East	25.80	21.13	12.26	10.89	29.91	100
South West	30.41	23.68	11.58	9.51	24.82	100
Wales	30.48	9.62	33.52	6.09	20.29	100
Scotland	79.69	3.02	12.30	1.31	3.69	100
Northern Ireland	83.97	3.02	4.18	4.97	3.82	100
Aggregate UK	**42.16**	**18.03**	**12.34**	**7.53**	**19.94**	**100**

Note: The sum of the five migration types for each region equals to 100%.

significantly across regions, with Scotland having the highest proportion of non-Mover at 83.97%, consistent with previous studies. Conversely, the East of England shows a low non-mover percentage of 13.99%, possibly due to shifts in employment sectors and reliance on lower-skilled EU workers.

Regions with high pay and household wealth, such as South East and South West, see significant migration to London for employment, driven by better job prospects and higher salaries. London attracts graduates primarily from the East of England, South West, and South East, comprising 61.49% of its employment from these regions.

Scotland's unique policy of free education and high retention rates within the region result in a high percentage of non-Mover and low move-returner. Similarly, Northern Ireland also exhibits high non-mover rates, while Wales struggles to retain graduates for employment, experiencing a 'brain drain' effect.

Stay-Mover and move-stayer represent a relatively small proportion of migration patterns, with London attracting a large net inflow of graduates due to its ability to retain and attract talent. East of England stands out with a high percentage of non-returning double Mover, indicating significant outflows of graduates after completing their studies.

The disparities in migration patterns between regions are influenced by geographical factors, regional wealth, university status, and policies, highlighting the complex interplay of factors shaping graduate migration behaviour.

2.4. Methodology

2.4.1. Defining migration

The majority of existing literature on graduate migration typically focuses on a singular aspect of mobility, such as movement either from domicile to Higher Education Institutions (HEIs) or from HEIs to employment. However, the research adopts a comprehensive approach by examining both stages of mobility simultaneously, similar to the approach taken by Kidd, O'Leary, and Sloane (2017). This methodology allows to provide a more complete representation of individuals' journeys from their pre-higher education domicile to early employment, which is arguably one of the most important occupation investments.

In this study, migration is defined using a distance-based measure. Essentially, an individual is classified as a migrant if the distance travelled exceeds a specified threshold. Conceptually, migrants in this study are those who have relocated from their domicile in the first stage and/or from the location of education in the second stage. However, given the input of postcode and distance by the graduate, it is imperative to establish a distance threshold that ensures a move significant enough to qualify as a departure from the domicile and educational region.

Determining the appropriate distance threshold is an empirical question. Previous studies, such as those by Gordon (1991), Boyle (1993), and Gordon and Molho (1998), have shown that varying distance thresholds can offer insights into different migration patterns. For instance, Gordon's research suggests that individuals migrating distances of 25 miles or less often do so for housing-related reasons, such as financial constraints. Therefore, if the distance between domicile and Higher Education Institutions (HEIs) and/or between HEIs and employment is less than 25 miles, it is probable that the individual remains in the same location, as indicated by previous studies.

Additionally, Gordon suggests that distances greater than 100 miles are more likely to be associated with employment-related migration. However, it's worth noting that Gordon's study primarily focused on the migration of individuals who were already employed, potentially leading to an overestimation of the 100-mile threshold. Nevertheless, it is intriguing to explore whether migration distances vary across the five different migration categories considered in this study.

However, unlike Gordon, this study predominately focuses on the migration choices of graduates while entering their higher education studies as well as their migration choices while searching employment after graduation. Therefore, in comparison with this research and the literatures previously considered, the level of heterogeneity in terms of objective is not significant. The study by Faggian and McCann (2008) considered migration as a movement covering a distance of more than 15km. The study by Faggian et al investigated the interrelationships between the interregional flows of human capital, and innovation migration behaviour of well-educated UK university graduates from university into employment. The findings are then examined to relate the human capital flows to both the labour market characteristics as well as the knowledge characteristics of the employment region. Since the study by Faggian et al. (2008) is comparatively consistent with the research, the distance threshold of 15km is adopted. This distance threshold is suitable as it largely captures local HEIs and local labour market regions in the UK.

However, unlike Gordon's focus on employed individuals, this study primarily examines the migration decisions of graduates as they enter higher education and seek employment after graduation. Thus, compared to previous literature, there is not a significant difference in terms of objectives. For instance, Faggian and McCann (2008) defined migration as movement covering a distance of more than 15km. Similarly, Faggian et al. investigated the interrelationships between interregional flows of human capital and the migration behaviour of well-educated UK university graduates from university to employment. They then related these findings to labour market and knowledge characteristics of the employment region.

Utilising Travel to Work Areas (TTWAs) is another method to determine an appropriate radius distance. However, research by Faggian et al. (2013) suggests an average radius in the UK of approximately 15.68km, closely resembling the 15km threshold. Thus, both TTWA results and the 15km boundary show similar migration patterns. To increase the credibility of the findings, robustness tests are conducted using alternative distance thresholds of 25km, 50km, and 100km. These thresholds are chosen empirically based on migration patterns, as explained in detail in the descriptive statistics section.

Given the consistency of the research with the study by Faggian et al. (2008) and Faggian et al. (2013), a distance threshold of 15km is adopted. Also, using this distance threshold allows for comparability with prior studies on graduate migration in the UK, including those by Faggian and McCann (2006; 2009), Faggian et al. (2006), and Abreu, Faggian, and McCann (2014). This threshold effectively captures local Higher Education Institutions (HEIs) and labour market regions in the UK. Faggian et al. applied a 15km distance threshold, as nearly all urban labour markets in the UK fall within this range. The primary exception to this is London, being the largest and most densely populated region in the UK, which naturally has a larger urban labour market radius compared to other regions. However, from a travel time perspective, London is often viewed as comprising several distinct urban market areas, each with a radius of less than 15km.

However, there is a slight but major difference in the method of analysing the distances graduate's travel. Most studies focus on graduates that migrate, although, this research incorporates the category of 'Non-Mover' as the base. Using the 15km as a threshold, the 'Non-Mover' category reflects the graduates that migrates less than 15km for both the first and second stage. This migration group is utilised as a base to compare the other graduate movement categories i.e. either or both the migration stages are greater than 15km. Including the 'Non-Mover' category, there will be five different migration categories, this will be further explained in the methodology section.

2.4.2. Model specification

Our analysis aims to uncover hidden factors influencing individuals' decisions to migrate between domicile, Higher Education Institutions (HEI), and employment. Following a two-stage modelling process similar to Kidd, O'Leary, and Sloane (2017), the investigation explores whether more mobile graduates earn a wage premium due to enhanced job opportunities. The study focuses exclusively on undergraduates, allowing to delve into how personal characteristics and human capital shape migration decisions across both stages.

Consider individuals are denoted as i and the two stages of migration is denoted as s whereby s takes the value of 1 or 2 for the first or second stage of migration respectively. Therefore, the indicator variable takes the form of I_{is}. For example, the first stage for individual i is depicted as the indicator I_{i1}. If individual i chooses to migrate at the first stage, the individual is classified as a mover for the first stage and takes a value of 1. Conversely, if the individual does not migrate at the first stage, the individual takes a value of 0 at the first stage of migration.

Furthermore, the second stage for individual i is depicted as the indicator I_{i2}. If individual i chooses to migrate at the second stage, the individual is classified as a mover for that second stage and takes a value of 1. Conversely, if the individual does not migrate at the second stage, the individual takes a value of 0 at the second stage of migration. This is shown below:

$$I_{is} = \begin{cases} 1 \text{ if migrate} & (\text{for } s = 1,2) \\ 0 \text{ if not migrate} & \end{cases} \tag{2.1}$$

To comprehend an individual's migration decision-making process, a simple probability model can be created for the first stage of this movement. Here, the student decides whether to study in their domicile region or migrate elsewhere. Let's assume there are *u* number of potential university locations the individual can choose from, and various factors directly influence this distance deterrence choice. As a basic probability model, let's consider that these direct factors encompass personal and family characteristics, impacting the individual's decision-making regarding distance deterrence (Gordon, 1978). Therefore, the potential utility function of the i^{th} individual movement from location D to u can be illustrated as follows:

$$U_{iu}^{*} = U_{iu}^{*}\,(\mathbf{A}_i, p_{iu}, f_{Du}) \tag{2.2}$$

where $\mathbf{A}_i$ denotes a vector of personal human capital characteristics, while p_{iu} denotes the expected returns to higher education from the individual i at the chosen location u. Finally, f_{Du} represents the distance deterrence factor between the individual's domicile location D and the potential university location U. Specifically, it captures the influence of distance-related factors on the decision to migrate for higher education.

Therefore, utility in this scenario represents the individual’s expected returns to human capital in dependence of factors such as the choice of degree, university and the degree class achieved. These features determine the individuals long term expected utility, directly impacting their employment prospects. Consequently, the current utility measure can be used, (U_{iu}^{*}) while including the direct factors, $(\mathbf{A}_i, p_{iu}, f_{Du})$ to assume the utility comprises a deterministic segment, $[S\,(\mathbf{X}_i, \mathbf{Y}_u, f_{Du})]$. This segment is linear with respect to the potential utility function and the error term e_{mu} and has a generalized extreme value distribution (Faggian, McCann and Sheppard, 2007). Therefore, equation (2.2), can be expressed as an additive random utility model:

$$U_{iu}^{*} = \boldsymbol{V}_{iu}\,(\mathbf{A}_i, p_{iu}, f_{Du}) + e_{iu} \tag{2.3}$$

Equations (2.2) and (2.3) can be incorporated into the concept of the main model methodology to understand how choices are determined, why some individuals decide to migrate when their counterparts do not, and if they migrate to attend university, what factors influence their decision on how far to travel from their domicile.

In the second stage of the model, data from the DLHE allows to compute the straight-line distances between the three crucial locations: domicile, university, and employment. This will be accomplished using the Pythagoras Theorem to calculate distances. This analysis helps determine whether a graduate entering employment remains in the region where they attended university or migrates elsewhere. When examining the data, attention should be paid to individual personal characteristics and human capital. Personal characteristics, including age, ethnicity, gender, degree studied, distance travelled from domicile to university, and highest degree grade achieved, are linked to a graduate's human capital. We aim to capture the impact of different levels of human capital before attending higher education by including the variable "Higher Education provider Mission Group." This variable evaluates the quality of institutions (HEIs) based on assessments conducted by independent and governmental bodies.

To illustrate the migration between university and employment, denote this movement with indicator variable, I_{is}[7]. The graduate will exhibit the value of 1[8] if employed in a location different to the University studied and 0 otherwise. Hence, this model can demonstrate the propensity of individual i to migrate for both movements, from domicile to University and from University to employment as shown below:

$$I^{*}_{i1t} = \omega_{it}\delta + \varepsilon_{i1t} \tag{2.4}$$

$$I^{*}_{i2t} = \nu_{it}\gamma + \varepsilon_{i2t} \tag{2.5}$$

where for student i I^{*}_{i1t} and I^{*}_{i2t} are unobserved latent variables at time t. Assume ε_{1t} and ε_{2t} are normally distributed random error terms at time t. Denote the relationship between

[7] The relationship between U^{*}_{i} (the utility function) and I^{*}_{i} (the indicator variable for migration) is where individuals assess their expected utility when making migration decisions. An individual compares the expected utilities associated with different locations and subsequently migrate if the expected utility of migrating exceeds the utility of staying.

[8] If an individual migrates to their region of domicile for occupational purposes, the graduate is assigned a value of 1^{*}, indicating they are a returner.

observed and unobserved mobility status to be 1 if $I_{idt}^{*} > 0$ and 0 if $I_{idt}^{*} = 0$ respectively (where $d = 1, 2$). The vectors ω and ν are the individual's characteristics, which influences the student's s choice to migrate, and γ indicates conformable vectors of returns to these characteristics. In theory, there is a possibility that the error terms in equations (2.4) and (2.4) are correlated, therefore in this research, a bivariate probit is employed to simultaneously estimate whether $\text{Cov}\ (\varepsilon_1, \varepsilon_2) = \beta$ holds true.

Given the prerequisite of the two different stages – first stage mobility: domicile to university, and the second stage mobility: university to employment, this thesis will imitate the classification implemented by DaVanzo (1983). Note that in the second migration stage, denoted as I_2, individuals who migrate within their region of domicile for occupational purposes (referred to as returners) are represented by 1^*.

- Non-Mover (Category 1): Graduates who study university and work in the same location as their home domicile ($I_1 = 0, I_2 = 0$)
- Move-Returner (Category 2): Graduates who migrate from their domicile location to attend university, although migrate back to their domicile location to work ($I_1 = 1, I_2 = 1^*$)
- Stay-Leaver (Category 3): Graduates who remain in their domicile location to attend university, although migrate elsewhere for employment ($I_1 = 0, I_2 = 1$)
- Leave-Stayer (Category 4): Graduates who migrate from their domicile location to attend university and stay in the same location for employment($I_1 = 1, I_2 = 0$)
- Non-Returning Double-Mover (Category 5): Graduates who migrate away from their domicile location to study university and migrate again for employment, although the graduate does not migrate back to their domicile location. This is different to the Move-Returner category as the second stage movement to employment is a location other than home domicile ($I_1 = 1, I_2 = 1^*$)

The utility function U_{iu}^{*} essentially drives the decision of the first migration decision represented by I_{i1t}^{*}. If U_{iu}^{*} is sufficiently high, the individual i is more likely to migrate for university, leading I_{i1t}^{*} to take a value of 1. If the utility of staying in the domicile region is higher, then I_{i1t}^{*} takes a value of 0. For the second stage of migration, If U_{iu}^{*} is sufficiently high, the individual i is more likely to migrate for their workplace, leading I_{i2t}^{*} to take a value of 1. If the utility of staying in the university region is higher, then I_{i1t}^{*} takes a value of 0. Therefore, there is an overall positive relationship between U^{*} and I^{*}.

In the primary regression analysis, samples meeting all three of the following criteria: permanent domicile, study, and employment in London, were excluded, known as "London Non-Mover" in the data. "London Non-Mover" constitute a subset of the broader category of "Non-Mover." They were excluded from the main regression because households in London are generally more affluent, leading to higher-quality schooling prior to university attendance and an increased likelihood of studying and working in London due to geographical proximity and the wage premium associated with the city. The inclusion of London Non-Mover could artificially inflate overall wage estimates due to the unique characteristics of London. This overestimation isn't due to a lack of graduate migration but rather because London's economy dominates graduate employment opportunities, a phenomenon known as agglomeration (Dunford and Fielding, 1992; McCann and Sheppard, 2001).

In the analysis of migration between domicile, university, and employment locations, there is flexibility to adjust the distance threshold according to preferences to differentiate between Mover and non-Mover. Following the approach of Kidd, O'Leary, and Sloane (2017), migration patterns and wage premiums are examined across distance thresholds ranging from 5km to 100km. The distance threshold increases incrementally by 5km, but the focus is specifically on four distinct thresholds: 15km, 25km, 50km, and 100km. This approach allows to investigate whether adjustments to the distance threshold led to changes in migration patterns and wage premiums, and if so, the extent of these changes.

There are two groups which are analysed in this sample represented as Mover and non-Mover. Given the two groups, group A is consequently denoted as Non-Mover (N), and group B as Mover (M). There is only one classification for group A the Non-Mover, labelled as category 1. Whereas, for group B, the Mover consists of four unique migration classification; category 2, 3, 4 and 5 as Move-Returner, Stay-Leaver, Leave-Stayer and Non-Returning Double-Mover respectively. By differentiating the four Mover categories, the wage premiums can be independently analysed with comparison to the Non-Mover.

2.4.2.1. Conditional distribution function

Previous studies, like Faggian, McCann, and Sheppard (2007), primarily focus on standard linear regression techniques, estimating the average relationship between the outcome variable (earnings) and independent variables (such as migration) using methods like normal OLS regressions and logit models. However, these traditional models only provide insights into the overall conditional mean and do not capture the relationship between different points in the earnings distribution. To address this limitation, conditional quantile regression techniques (Koenker and Bassett, 1978) within a correlated random effects framework are utilised. This approach allows to analyse earnings results across the entire conditional distribution, providing a more comprehensive understanding of the relationship. The estimated conditional quantile (Φ^{th}) earnings can be written as:

$$Q^{\Phi}\ (\bar{y}^{m} \mid x^{m}) = \ x^{m}\beta^{\Phi,m} \tag{2.6}$$

where Q denotes the estimation with the respect to the varying quantile Φ. $\bar{y}^{m}$ are the log earnings for a specific mobility status group m, where m associates with the categories from 1 to 5 as defined earlier. x is a vector that consists of the personal characteristics of individuals which influence earnings, given that vector x incorporates each of the migration categories, though the quantile distribution is adjusted across each category. $\beta^{\Phi,m}$ are the outcome of coefficient returns produced by conformable vector between the contrasting characteristics with specific quantile Φ and migration group m. By using the conditional quantile regression approach, it enables to estimate the earnings for Mover and Non-Mover throughout the entire quantile of the distribution.

2.4.2.2. Unconditional distribution function

Combining quantile regression with the conditional distribution function (CDF) offers valuable insights beyond just the mean of a dependent variable. It allows for a deeper understanding of the distribution, forming the basis of estimation. However, a limitation is the lack of focus on the effects for quantiles of the marginal distribution of the dependent variable Y (I^{*}_{imt} for this research). Therefore, the unconditional distribution function is introduced. This method allows the evaluation of the impact of changes in the distribution of the explanatory variables (e.g., subject studied) on quantiles of the unconditional distribution of the dependent variable, i.e., earnings.

To further investigate the distributional relationship between graduate migration and earnings, a method similar to Ferrer-i-Carbonell and Frijters (2004) is integrated - unconditional quantile regression within a correlated random effects framework. Given the data's discrete nature, a data smoothing technique is applied, which is a necessary adjustment for the issue at hand (Machado and Silva, 2005). This is essential because the outcome variable Y, representing earnings, exhibits a discrete distribution when estimated under quantile regression, as illustrated below:

$$Q_Y(a \mid \mathbf{X}) \tag{2.7}$$

where Y is the outcome variable, $\mathbf{X}$ is random variables and let equation (2.7) denote the $100\alpha^{\text{th}}$ quantile of the conditional distributions. It is not possible for equation (2.7) to be a continuous function of the parameters of interest due to the issue of the discrete distribution.

The problem can be resolved by constructing a continuous random variable for individual Y at time t whose quantiles have a known relationship with the quantiles of reported earnings, Y_{it}. This is achieved by generating an additional variable Z_{it} given that:

$$Z_{it} = Y_{it} + u_{it} \tag{2.8}$$

where u_{it} represents a random variable, which is independent of Y and $\mathbf{X}$ in equation (2.7) is distributed homogeneously within the interval [0, 1), this is also known as jitter the data

(Machado and Silva, 2005). Using quantile regression while implementing jittered data allows the distribution function of Z_{it} to be continuous, thus enables the regression method to be utilised to a monotonically transformed function of Z_{it}.

2.4.2.3. Oaxaca-Blinder – Decomposition of average earnings differentials

In this study, two main regression analyses are conducted. The first stage estimation involves the Oaxaca-Blinder (OB) decomposition, while the second stage estimation utilises the technique of "recentred influence function," an advancement of OB decomposition. The OB decomposition, pioneered by Oaxaca and Blinder in 1973, offers a deeper understanding compared to traditional OLS regression by breaking down different components of the average result. It produces three distinct outcomes: the overall effect, wage structure effect, and composition effect. The sum of the wage structure effect and the composition effect equals the overall effect. We employ this regression technique because traditional OLS regressions often misidentify the error term as wage discrimination, neglecting other factors that may contribute to wage differences.

Determining which explanatory variables to include or exclude in any regression poses a challenge for econometricians and may lead to errors, such as omitting valid explanatory variables, rendering their effects unobserved. This issue is exemplified in the study by Card and Krueger (1992), where they explored school quality differences between black and white citizens in the US and their relative earnings. While many papers attributed the wage gap to poorer education quality for black citizens, Card and Krueger argued that discrimination also played a significant role in the wage discrepancy.

The OB decomposition employed shares similarities with the method used by Fortin, Lemieux, and Firpo (2011). It's designed to extend the decomposition to distributional parameters beyond the mean, such as the Gini coefficient and quantiles. The OB is often seen as a straightforward accounting estimate based on correlations, akin to an OLS estimate. In essence, the OB operates much like an OLS, except that it's estimated by substituting the parameter vector, β_t with the OLS estimates, as well as replacing the expected value of the covariates of the conditional expectation $\mathrm{E}[\mathrm{X}\ |T = t]$ by the sample averages (Firpo, Fortin and Lemieux, 2018).

Incorporating a quantile regression model allows to analyse various distribution levels beyond just the mean of the wage distribution, unlike standard OLS estimates. By examining percentiles like the 10th, 25th, 50th, 75th, and 90th, a comprehensive understanding of the wage distribution is gained. Previous studies, such as O'Leary, Murphy, and Blackaby (2004), have also utilised these percentiles for comparison.

Integrating the standard OB decomposition by Oaxaca and Blinder (1973) with specific percentiles enables to assess differences in wage structures between two groups of graduates (A and B) across the distribution spectrum. Group A consists of Non-Mover, while Group B comprises Mover. We'll use these categories for simplicity in illustrating the Oaxaca-Blinder method, which estimates the difference in mean outcomes between Groups A and B. Since the OB decomposition is similar to an OLS regression in many respects, the main assumption remains consistent. For instance, the main assumption of the OB is that the outcome variable Y (earnings) is linearly correlated to the covariates, X as well as the residual term u is conditionally independent of X such that:

$$Y_{it} = \beta_{i0} + \sum_{n=0}^{N} X_{cn}\beta_{in} + u_{ic} \tag{2.9}$$

where, i = group A or B, $\mathrm{E}\left(u_{gc}\middle|\, X_c\right) = 0$ and X is the vector of covariates ($X_c = [X_{c1,\dots,}X_{cn}]$). We define equation (2.9) as the mean for either group A or group B. By subtracting the means, the overall difference of the averages of earnings due to the consequence of graduate migration can be found. Y_{ic} is the aggregate of both the composition and wage structure effect, which gives the following equation:

$$\tilde{\omega}_O^{\mu} = \hat{\Upsilon}_B - \hat{\Upsilon}_A \tag{2.10}$$

where $\tilde{\omega}_O^{\mu}$ represents the overall difference, $\hat{Y}_B$ is the average outcome of group B (treatment group i.e. Mover) and $\hat{Y}_A$ is the average outcome of group A (control group i.e. Non-Mover). Therefore, if there is a positive (negative) wage premium with respect to graduate migration, then then overall effect, $\tilde{\omega}_O^{\mu}$ will give a positive (negative) earning result.

Using equation (2.10), the analysis of earnings gaps through a decomposition framework allows for the breakdown of mean earnings differentials into two distinct effects: the composition effect and the wage structure effect (Firpo et al., 2009). In this method, known as the Oaxaca-Blinder decomposition, the focus is primarily on the wage structure effect, which represents the pay premium associated with the variable of interest between the treatment and control groups.

In this study, the pay premium reflects the rewards individuals receive for migrating (either in the first or/and second stage) compared to their non-migrating counterparts. To illustrate both effects, equation (2.10) can be decomposed as follows:

$$\tilde{\omega}_{O,Q\theta} = \underbrace{\left[\widehat{X}_1\,\Omega_{1,\theta} - \widehat{X}_2^C\,\Omega_\theta^C\right]}_{\text{(Unexplained)}} + \underbrace{\left[\widehat{X}_2^C\,\Omega_\theta^C - \widehat{X}_2^C\,\Omega_{0,\theta}\right]}_{\text{(Explained)}} = \tilde{\omega}_{W,Q\theta} + \tilde{\omega}_{C,Q\theta} \qquad (2.11)$$

where $\widehat{X}_1$ and $\widehat{X}_2$ is the sample means of earnings for the graduates that migrated either the first or/and second stage movement and graduates that did not migrate in either the movement stages (i.e. group A and group B), respectively. $\widehat{X}_2^C$ is the sample mean of earnings for the counterfactual distribution at chosen percentiles. Finally, the θ term indicates the chosen percentile which is the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th}. Equation (2.11) is the overall effect; it has two separate constituents. The first term of equation (2.11) is the "unexplained" component of the OB decomposition. As the OB in this research mainly focuses on wage differentials due to sample discrimination, this first term will be referred to as, unexplained effect as the "wage structure" effect ($\tilde{\omega}_{W,Q\theta}$), similar to Fortin, Lemieux and Firpo (2010). While the second term of the OB, $\tilde{\omega}_{C,O\theta}$ is the "explained" component of the estimate. The explained component is the composition effect of the OB decomposition where it captures the differences in covariates.

2.4.2.4. Limitations and resolve of Oaxaca-Blinder decomposition

The OB decomposition technique has two main limitations. Firstly, it relies on the assumption of linearity in the conditional expectation, similar to ordinary least squares (OLS) estimates (Barsky et al., 2002). To address this linearity issue, a non-parametric technique can be employed to estimate the conditional expectation $\mathrm{E}[\mathrm{X} \mid T = t]$ under a non-parametric technique. Barsky suggested using a non-parametric reweighting approach, as proposed by DiNardo (2002), to estimate the decomposition. While this non-parametric method offers simplicity and can be applied under more general distributional statistics, it does not directly allow for further decomposition of each covariate concerning the wage structure and composition effect.

Over time, advancements in the OB technique have enabled additional decompositions of the two effects without compromising estimation consistency. For instance, DiNardo (2002) introduced a method that computes the overall wage structure and composition effects for various distributional statistics. This was further enhanced by Fortin, Lemieux, and Firpo (2018), who developed a reweighting procedure to divide the two overall effects (wage structure and composition effects). This decomposition distinguishes the contribution of each covariate, allowing for a more detailed analysis, as discussed in the following section.

2.4.2.5. Overview of the first stage estimation

In the initial analysis, the Oaxaca-Blinder decomposition method is utilized to split the earnings distribution into two components: the wage structure effect and the composition effect. This involves applying reweighting functions, $\Omega_1(T), \Omega_0(T)$ and $\Omega_c(TX)$ which can be estimated parametrically or non-parametrically the first two reweighting functions adjust characteristics of the wage structure distribution, W_1 if $T = 1$ and W_0 if $T = 0$. The third reweighting function converts the characteristics of the marginal distribution of W into the characteristics of the counterfactual distribution of W_0 in the situation of $T = 1$. Therefore, the counterfactual function is typically a vector which is constructed artificially using the observed features in the population of the base group as the foundation.

Finally, the distributional statistics ν_1, ν_0 and ν_c are estimated by using the reweighted samples. ν_c is the counterfactual distributional statistic and the difference between the two groups of ν's i.e. ν_0 and ν_1, results to the ν-overall wage gap. In simple terms, the ν - overall wage gap is the differences in earnings measured with respect to the distributional statistic, ν.

Therefore, by using this developed method of the standard OB decomposition, it enables the details of the overall effect from the wage structure and composition effects to explain the pay differences between non-migrating graduates and migrating graduates at chosen points in the pay distribution (FFL, 2007 and 2018). The regression of this estimate is the result of equation (2.10):

Equation (2.10) illustrates the coefficient estimates for the graduates that migrate and those that do not at each percentile. $\tilde{\omega}_{O,Q\theta}$ represents the overall effect of the differences in earnings between the control group (Non-Migrating Graduates) and the treatment groups (4 category of Migrating Graduates). The overall effect is generated by two constituents, the unexplained component $\left(\tilde{\omega}_{W,Q\theta}\right)$and the explained component $(\tilde{\omega}_{C,Q\theta})$ which is the wage structure and composition effect respectively.

2.4.2.6. Detail of the wage structure effect and composition effect

Following from equation (2.11), group B is noticeably the treatment effect as it represents graduates that migrates, whereas group B represents the graduates that does not migrate. The raw wage gap (premium), ω can be decomposed as the sum "effect" of migration on graduate migrants, and the composition effect, ω_c is linked to differences in covariates between migrants and non-migrants' graduate. To summarise, the effect of migration for each graduate can be considered $(Y_{Bi} - Y_{Ai})$ as the individual treatment effect, while ω_W is the "average treatment effect on the treated" (ATT). Therefore, by the constitution of the equations, the composition effect can be evaluated from the estimated treatment effect as; ATT = $\tilde{\omega}_{W,Q\theta}$ and $\tilde{\omega}_{C,Q\theta}$ = $\tilde{\omega}_{O,Q\theta}$ - $\tilde{\omega}_{W,Q\theta}$. This is just the rearranging of equation (2.11) where the overall effect, $\tilde{\omega}_{O,Q\theta}$ is the aggregate of the wage structure effect, $\tilde{\omega}_{W,Q\theta}$ and the composition effect, $\tilde{\omega}_{S,Q\theta}$.

There are two main assumptions of the wage structure and composition effect: Assumption of ignorability and Assumption of Overlapping Support. The assumption of ignorability enables the identification of the effect of the treatment on the treated group (B). Let (T, X, u) have a joint distribution. Therefore, the distribution of the unobserved independent variables in the wage estimate is identical for group A and B when condition the vector of observed component. Whereas the assumption of overlapping support entails that the is an overlap in observable characteristics throughout the two groups, where there is no value for a vector of covariates i.e. x in X which solely observes among the graduates in group B. Therefore, for all x in X, $p(x) = \Pr\,[T = 1 \mid X = x] < 1$.

Under there two assumptions, the identification of the wage structure and composition effects is as follows:

- $\omega_{W,Q\theta}$ and $\omega_{C,Q\theta}$ are identifiable from data with (T, X, u)
- If the groups are identical i.e. Group A (., .) = Group B (., .) $\therefore\ \omega_{W,Q\theta} = 0$
- If the counterfactual distributions are identical i.e. $\mathrm{F}_{X|T=B} = \mathrm{F}_{X|T=A}$ then $\omega_{C,Q\theta} = 0$

The above provides the results of the OB decomposition under the two assumptions. The identification of $\omega_{W,Q\theta}$ and $\omega_{C,Q\theta}$ is captured from the quantities from the distributions that are obtained by integrating weighting for the sample with the inverse probabilities subject of either belonging to group A or group B given T. The second result points out that in the circumstance when there are no group differences in the wage estimation, then it is predicted there will be no wage structure effects as the treatment group will not have an earnings premium in comparison with the control group. The third result states that is there are no differences in the distribution of the covariates between group A and group B, then there will be no composition effects.

2.4.2.7. Further decomposition of the wage structure effect and composition effect

The conventional Oaxaca-Blinder (OB) decomposition method devised by Oaxaca and Blinder in 1973 will be adapted to allow for a more detailed breakdown of the wage structure effect and composition effect that have already been decomposed. This refined decomposition process, known as the "second stage estimation" (FFL, 2018), employs the recentred influence function (RIF) technique.

In the first stage estimation, the assessment focuses on whether there exists a wage premium for graduates who migrate in either or both the first and second stages. This is investigated by applying an OB decomposition to dissect the explained and unexplained components of the estimate, namely the composition and wage structure effects of the overall effect.

The second stage estimation extends beyond the standard OB approach by further decomposing the overall effect into its explained and unexplained parts. This enables to break down each of these constituents into a composition effect and wage structure effect. The objective is to provide a more detailed analysis by delineating the composition effect and wage structure effect within both the overall effect and the structure effect.

To achieve this, it is vital to decompose the wage structure and composition effects into their specific contribution of each individual covariate. To enable this, the method developed by Fortin, Lemieux, and Firpo (2010) is utilized., to compute the partial effects of changes in the distribution of covariates on a specified component of the distribution $\Upsilon_t | T$. The technique imposed by FFL (2010) uses a linear approximation to estimate a non-linear function of the distribution. The linearization technique enables to compute approximate partial effects of changes in the distribution of every covariate on the observed function. This method of further decomposition will be complemented by examining various distribution points, such as the 10th, 25th, 50th, 75th, and 90th percentiles. Analysing wage distribution is essential because relying solely on the sample mean can be overly simplistic (Disney and Gosling, 1998), as it fails to adequately capture the characteristics of both migrating and non-migrating graduates.

2.4.2.8. Counterfactual and the wage structure setting

In constructing a decomposition of the overall difference, both the wage structure and composition effects heavily rely on how the counterfactual wage distribution is formed. The purpose of introducing a counterfactual component is to model what would have happened if the wage distribution of graduates had a different characteristic from the actual sample. The counterfactual assumes that graduates in Group A are paid as if they were in Group B. This allows to investigate the impact of Mover (Group B), while considering the characteristics of Non-Mover (Group A) and keeping all other components of the wage structure constant within the Mover sample. Thus, the counterfactual aims to adjust the wage structure framework to reflect both observed and unobserved characteristics of graduates in different groups. Below are the two functions encapsulated by the wage structure:

Structural Form: This assumption states that the wage structure values between group A and group B can only be different from three reasons. To see this, let assume a graduate i is either from group A or B and is earning with the wage structure, w_A and w_B respectively. Given the graduates' explanatory variables (X) and unobservable (ε), the graduate function can be derived as:

$$Y_{Ai} = W_A(X_i, X_{\varepsilon i}) \text{ and } Y_{Bi} = W_B(X_i, X_{\varepsilon i}) \tag{2.12}$$

We can see the three potential causes for the variations with equation (2.12); Firstly, the wage setting functions are different, W_A and W_B; Secondly, the distribution of explanatory variables (X) is different; Thirdly, the unobservable (ε) is different. Equation (2.12) is composed of three different components, thus, when there is a dissimilarity between group/individual A and B with any of the three components, the wage structure will be different. The objective of the overall decomposition is to separate and isolate the component of the first factor i.e. differences between W_A and W_B, from the observable and unobservable variables in order to examine each constituent of the decomposition.

Therefore, the proposed counterfactual chosen for the treatment group will alter the result of group B’s wage structure.

Simple Counterfactual Treatment: The counterfactual wage structure, m^C is referred as a simple counterfactual treatment when the following criteria are satisfied; either i) m^C (.,.) ≡ m_A (.,.) for graduates in group B, or ii) i) m^C (.,.) ≡ m_B (.,.) for graduates in group A, where ^{C} represents the counterfactual wage. This assumption exemplifies that an additional counterfactual wage structure i.e. m^* (.), illustrates the salary of a graduates if there was no migration.

These two examples demonstrate that in an OB decomposition, the counterfactual incorporated is vital as it determines the influence of the structural wage setting functions since it links the observed and unobserved variables of the graduates to their wages for group A and B i.e. migrators and non-migrators.

2.4.3.1. Recentred Influence Function (RIF)

RIF examines the impact of each covariate on a specific function of interest, an approach that is practical and capable of estimating how increasing the proportion of migrating graduates affects average salaries is needed, especially when i ($i = 0$: non-migrate, $i = 1$: Migrate) is not univariate and binary. Therefore, the influence function (IF) method is employed, widely used for assessing the robustness of statistical and econometric models. The IF, denoted as, $\mathrm{IF}(Z_{it}; v)$ measures an individual observation's influence on the distributional statistic v of interest. By integrating v into the IF, the recentred influence function (RIF), is derived, introduced by Firpo, Fortin, and Lemieux (2009).

The RIF is valuable because it ensures that the expectation of the regression matches the distributional statistic v. Given its flexibility in estimating various distributional statistics, the method by FFL (2009) can extend straightforwardly to alternative quantiles beyond just mean estimation, encompassing variance, Gini coefficient, and other measures of inequality.

The RIF regression enhances equation (2.10), which provides an aggregate decomposition of the wage structure and composition effects. This, combined with percentiles, allows to explain salary differences between migrating and non-migrating graduates across various distribution points. FFL's technique (2007 and 2018) facilitates estimating relationships between distributional statistics using RIF, capturing the marginal effect on unconditional quantile shifts in covariate distributions while holding other factors constant.

As explained before, a RIF regression is similar to a standard OLS regression, except that the outcome variable, Y (salary), is substituted by the statistic of the recentred influence function of interest. We will use the auxiliary variable Z_{it} from equation (2.8) i.e. $Z_{it} = Y_{it} + u_{it}$, to construct the RIF regression. Let $\mathrm{IF}(Z_{it}; v)$ be the influence function that represents the observed earning, Z_{it} for the distributional statistic of interest, $v\ (F_z)$. The transformed function is then incorporated with a supplementary model introduced by Firpo et al (2009). The recentred influence function framework substitutes the disturbance statistic of interest v, as a replacement of the dependent variable, Z_{it}. Therefore, the aggregation of the influence function and the distributional statistic of interest forms the RIF given as, $\mathrm{RIF}\ (Z_{it}; v) = v\ (F_z) + \mathrm{IF}(Z_{it}; v)$, thus rearranging shows that the statistic of interest is derived as given, $v\ (F_z) = \int \mathrm{RIF}\ (Z_{it}; v)\ .\ dF(Z_{it})$. This can be simplified where the unconditional expectation of the RIF $(Z_{it}; v)$ can be illustrated as a linear function of the independent variables as follows:

$$\mathrm{E}\ [\mathrm{RIF}\ (Z_{it}; v)\ |\ X] = X_{\Upsilon} + \varepsilon \qquad (2.13)$$

where the parameters of Υ can be estimated by OLS.

Since quantiles are being incorporated in this study, the influence function, $\mathrm{IF}\ (Z_{it}; v)$ is given by $\frac{(\theta - \Phi\ \{Z \le Q_\theta\})}{f_Z(Q_\theta)}$, where θ is the quantile, Q_θ is the population of the outcome variable Z and $\Phi\ \{\ \cdot\ \}$ is an indicator function and $f_Z\ (\ \cdot\ )$ is the measure of marginal distribution of Z. Therefore, $\mathrm{RIF}\ (Z_{it}; v)$ is equivalent to $Q_\theta + \mathrm{IF}\ (Z_{it}; v)$, and for the θ quantile the recentered influence function takes the following form:

$$\text{RIF}\ (Z_{it}; Q_\theta) = \ Q_\theta + \frac{\theta - \ \Phi\{Z \leq Q_\theta\}}{f_Z(Q_\theta)} \tag{2.14}$$

where Q_θ is the unconditional quantile of the observed earning, Z_{it} and the function f_Z (Q_θ) is the marginal distribution of Z_{it} at the quantile Q_θ. The RIF can be constructed by substituting the distributional statistic of interest by a given percentile, θ. So given the conditions, the RIF at the θ^{th} percentile is the following:

$$\text{RIF}\ (Z_{it}; Q_\theta) = \ Q_\theta + \ \text{IF}\ (Z_{it}; Q_\theta) \tag{2.15}$$

Equation (2.15) is equivalent to an unconditional quantile regression which is implemented by FFL (2009). The regression estimates using the RIF while introducing a correlated random effects model which attributes the characteristics of unobserved heterogeneity by using a fixed effect. This model gives the following:

$$\text{RIF}\ (Z_{it}; Q_\theta) = \ \alpha_i + \ x_{it}\beta + GM_i\Upsilon_1 + \ y_i\Upsilon_2 + \ \widehat{\text{X}}_i\sigma + \ \varepsilon_{it} \tag{2.16}$$

where x is a vector determined by time, t and explanatory variable β , $\widehat{\text{X}}$ denotes the proxy of individual fixed effect where it results the mean value of the time-varying explanatory variables over time, y is a vector of explanatory variables while keeping time constant, GM is a coordinated group of dummy variables that signifies the type of migration (given the base variable is Non-Mover i.e. graduate does not move in both stages of migration) and the model is estimated for a chosen specific quantile θ.

2.4.3.2. Decomposition terms from the second stage estimation (RIF)

This section outlines the potential outcomes derived from the equations of the RIF technique. The second stage estimation of this study gives an insight on the covariates of the first stage estimation i.e. standard OB decomposition. The further decomposition of the wage structure and composition effect (RIF) provides the four possible decompositions of the overall difference, $\tilde{\omega}_{O,Q\theta}$:

D.1: Differences related to the return to observable variables from the structural *m* functions. For instance, the counterfactual set up states the situation where all components are the same for group A and B except for the return X.

D.2: Differences related to the return to unobservable variables from the structural *m* functions. For instance, the counterfactual set up states the situation where all components are the same for group A and B except for the return ε.

D.3: Differences related to the distribution of the observable characteristics. For instance, the counterfactual set up states the situation where all components are the same for group A and B except for the distribution of X.

D.4: Differences related to the distribution of the unobservable characteristics. For instance, the counterfactual set up states the situation where all components are the same for group A and B except for the distribution of ε.

Since unobservable variables is part of the OB component, it is only possible to decompose $\tilde{\omega}_O^{\mu}$ into four decomposition functions when various assumptions are involved between the joint distribution of the observable and unobservable variables. In addition, if further separability assumptions are not implemented on the structural forms of the *m* functions, it is practically unfeasible to isolate the contribution of the returns between the observables and the unobservable variables. Therefore, when performing a standard OB decomposition regression in returns, it is very difficult to separate the contribution of the return to each covariate separately. This implies that, by estimating with a standard OB regression, the mean of the earnings from the sample can only be decomposed into two separate components: the wage structure effect and the composition effect. Therefore, further break down of detailed decomposition is problematic.

2.4.3.3. Overview of the second stage estimation

The second-stage estimation incorporates the RIF regression, as proposed by Firpo, Fortin, and Lemieux (2018). This allows for a further decomposition of the Oaxaca-Blinder (OB) decomposition from equation (10). Specifically, the RIF regression breaks down the wage structure effect $\tilde{\omega}_{W,Q\theta}$ and the composition effect $\tilde{\omega}_{C,Q\theta}$ to isolate the pure wage structure effect and the pure composition effect, while also identifying residual components that capture reweighting and specification errors.

To illustrate this, we begin by distinguishing the wage structure effect and composition effect based on equation (2.11):

$$\tilde{\omega}_{W,Q\theta} = \widehat{\mathrm{X}}_1 \left(\Omega_{1,\theta} - \Omega_{\theta}^{C}\right) + \left(\widehat{\mathrm{X}}_1^{C}\, \Omega_{\theta}^{C} - \widehat{\mathrm{X}}_2^{C}\right)\Omega_{2,\theta}^{C} = \tilde{\omega}_{W,pQ\theta} + \tilde{\omega}_{W,eQ\theta} \tag{2.17}$$

$$\tilde{\omega}_{C,Q\theta} = \left(\widehat{\mathrm{X}}_2^{C} - \widehat{\mathrm{X}}_2\right)\Omega_{2,\theta} + \widehat{\mathrm{X}}_2^{C}\left(\Omega_{\theta}^{C} - \Omega_{2,\theta}\right) = \tilde{\omega}_{C,pQ\theta} + \tilde{\omega}_{C,eQ\theta} \tag{2.18}$$

Equation (2.17) represents the wage structure effect, which consists of two parts: the pure wage structure effect $\tilde{\omega}_{W,pQ\theta}$ and the reweighting error $\tilde{\omega}_{W,eQ\theta}$. Similarly, equation (2.18) represents the composition effect, which also has two components: the pure composition effect $\tilde{\omega}_{C,pQ\theta}$ and the specification error $\tilde{\omega}_{C,eQ\theta}$.

Minimizing both the reweighting and specification errors in the RIF regression is crucial to ensuring that the model is well-specified and accurately estimates the wage structure and composition effects. Reweighting errors should be close to zero to indicate that the counterfactual reweighting is accurate, and specification errors should also be minimized to confirm that the model is properly specified, and linearity holds.

From equation (2.17), the contribution of individual covariates (*l* = 1, 2, …, *k*) of the pure wage structure effect and pure composition effect as follows:

$$\Delta_W(\tilde{\omega}_{W,pQ\theta}) = \sum\nolimits_{i=1}^{k} \Delta_W \ ((\widehat{\mathrm{X}}_{2,j}^{C} - \widehat{\mathrm{X}}_{2,j})\, \Omega_{2,j,\theta} \qquad (2.19)$$

$$\Delta_{C,pQ\theta} = \sum\nolimits_{i=1}^{k} ((\widehat{\mathrm{X}}_{2,j}^{C} - \widehat{\mathrm{X}}_{2,j})\, \Omega_{2,j,\theta} \qquad (2.20)$$

The assumption of linearity of the RIF regression is essential as this allows the changes in both the structure and composition effects to be evaluated with the changes in the overall pay gap at chosen specific percentile of the wage distribution at different time periods. The changes in the wage gap at two different time intervals can be illustrated as the following:

$$\begin{aligned} \Delta_W(\tilde{\omega}_{Q\theta}) &= \Delta_W \left(\left[\widehat{\mathrm{X}}_1\left(\Omega_{1,\theta} - \widehat{\mathrm{X}}_2^{C}\, \Omega_\theta^{C}\right]\right]\right) + \left(\left[\widehat{\mathrm{X}}_2^{C}\, \Omega_\theta^{C} - \widehat{\mathrm{X}}_2\, \Omega_{2,\theta}\right]\right) \qquad (2.21) \\ &= \Delta_W(\tilde{\omega}_{WQ\theta}) + \Delta_W(\tilde{\omega}_{CQ\theta}) \end{aligned}$$

$$\begin{aligned} \Delta_W(\tilde{\omega}_{WQ\theta}) &= \Delta_W \left(\widehat{\mathrm{X}}_1\left[\Omega_{1,\theta} - \Omega_\theta^{C}\right]\right) + \Delta_W \left(\left[\widehat{\mathrm{X}}_1 - \widehat{\mathrm{X}}_2^{C}\right] \Omega_\theta^{C}\right) \qquad (2.22) \\ &= \Delta_W(\tilde{\omega}_{W,pQ\theta}) + \Delta_W(\tilde{\omega}_{W,eQ\theta}) \end{aligned}$$

First, the Oaxaca-Blinder (OB) decomposition uses the sample from $T = 0$ as the baseline and compares it to a counterfactual sample where $T = 0$ characteristics are reweighted to resemble those at $T = 1$. This comparison allows us to isolate the pure composition effect, using the $T = 0$ sample as the reference for the wage structure.

Second, the decomposition is performed using the sample from $T = 1$ alongside the counterfactual sample, where the counterfactual sample is now used as the reference. This approach isolates the pure wage structure effect.

2.5. Results and discussion

All the results from this section are based on the investigation of the HESA where the data is gathered for the years between 2002 and 2012. The graduate migration and their earnings are examined yearly for each of the five migration groups while using the group "Non-Migrators" as the base for comparison. Moreover, the results of each year and groups will be analysed at different points of the wage distribution, with the chosen levels of the 10th, 25th, 50th, 75th and 90th percentiles which is the similar to the paper by Murphy, Blackaby, O'Leary and Staneva (2019). Finally, the regressions are based on the data from HESA where the earnings figure is pre-subjected to being weighted. Also, to primarily focus on how wages are affected by graduate mobility, all the results illustrated in this section will exclusively examine individuals that graduate from their first degree in HEI i.e. undergraduates.

2.5.1.1. Aggregate decomposition (First stage estimation)

The aggregate decomposition estimate given by equation (10) are presented in Figure 2.1. There are 4 separate groups for each of the "Mover" category where they are estimated by using the "Non-Mover" as the control. At each percentile interval, the graph shows the:

- Overall Mean Wage Gap ($\tilde{\omega}_{OQ\theta}$) – The total wage inequality between the "Mover" and "Non-Mover" as a result of each individual covariates.
- Overall Wage Structure effect ($\tilde{\omega}_{wQ\theta}$) – Refers to how wages differ between two groups while keeping both the groups characteristics constant. This effect accounts for how labour market conditions influence wage disparities for similar qualifications or experience. In this context, the overall wage structure effect measures the differences in premium due to whether an individual is a Mover or a Non-Mover. Therefore, the wage structure effect shows that wages are not only based on the characteristics (like education and experience), but also on how much those characteristics are valued in the specific job market or context where the person works. Therefore, Movers and Non-Movers with identical qualifications can earn

different wages because the "value" placed on their skills varies with location, industry, and market demand.

- Overall Composition effect ($\tilde{\omega}_{CQ\theta}$) – The overall composition effect (also known as the "characteristics effect") reflects the part of the wage gap attributable to differences in the characteristics of each group (for example, differences in levels of education and experience between the groups being compared). This effect considers how the distribution of these characteristics across groups contributes to wage disparities. For example, if Movers generally have higher levels of education or more specialised skills than Non-Movers, this would be part of the composition effect because it reflects differences in who is in each group.

2.5.1.2. Aggregate decomposition (Comparing between years)

The results for the aggregate decomposition are presented in Figure 1 – 4. Figure 1 represents the overall decomposition effect of the total change, composition and wage structure effect for each year. Figure 2 congregates the total change of each group onto a graph corresponding the year. Table 2.5 summarises the results for the standard measures of distribution (10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentile).

2.5.1.3. Analysis of the base year (2004/05)

In this section, the aggregate decomposition for the base year of 2004/05 will be analysed. The aggregate decomposition measures the overall effect, composition effect and the wage structure effect. Once the base year is analysed, comparison to the subsequent periods can be made. Primarily, Figure 2.1 and 2.2.2 to 2.2.4 will be examined in this section. Both these figures are derived from the aggregate decomposition which is reported on Table 2.5.1 to Table 2.5.4. It should be noted that the results from the decompositions are compared to the base group i.e. Non-Mover.

Figure 2.1

Overall effect for each mobility type (Yearly)

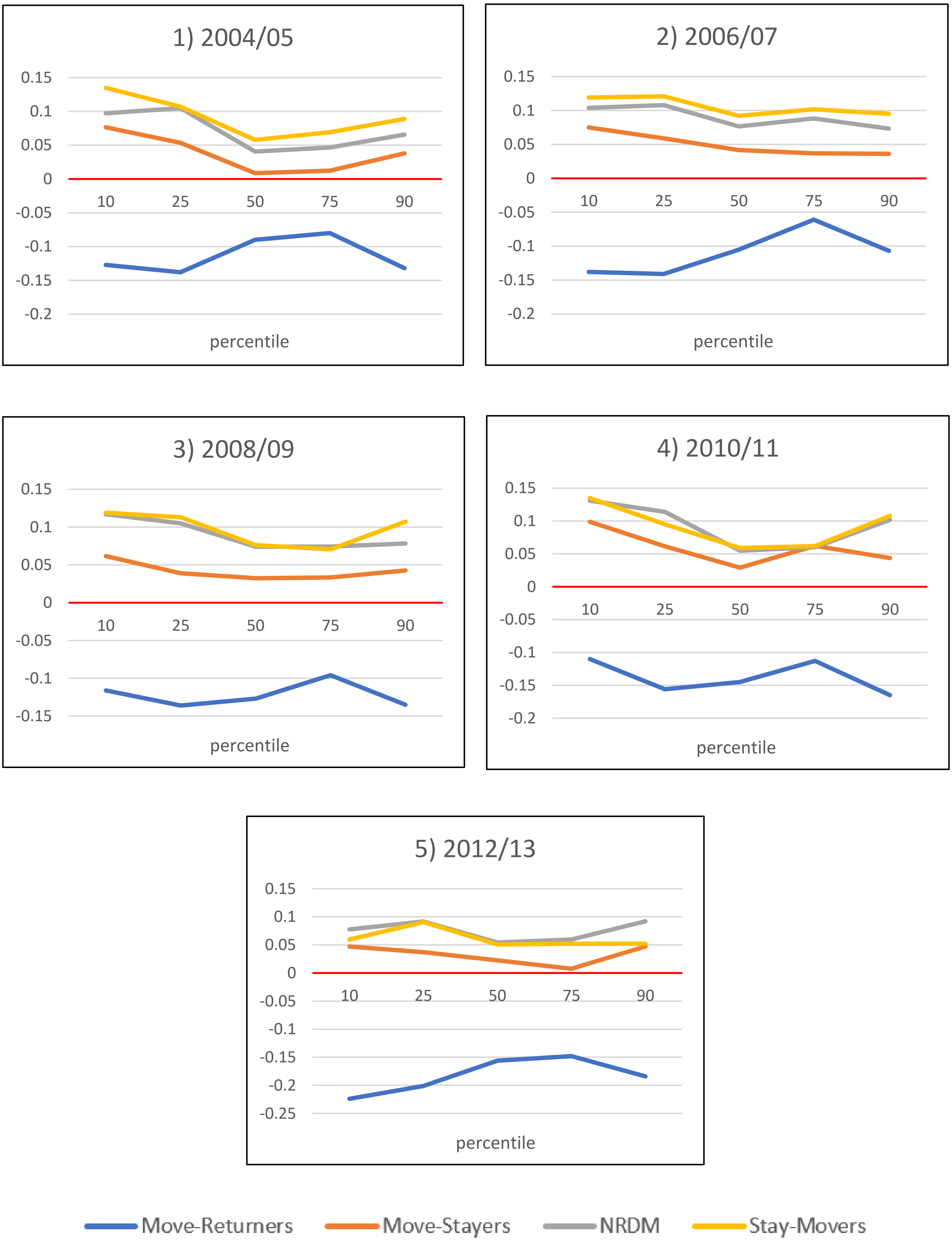


Figure 2.2.1

Overall decomposition effect for Move Returner (Yearly)

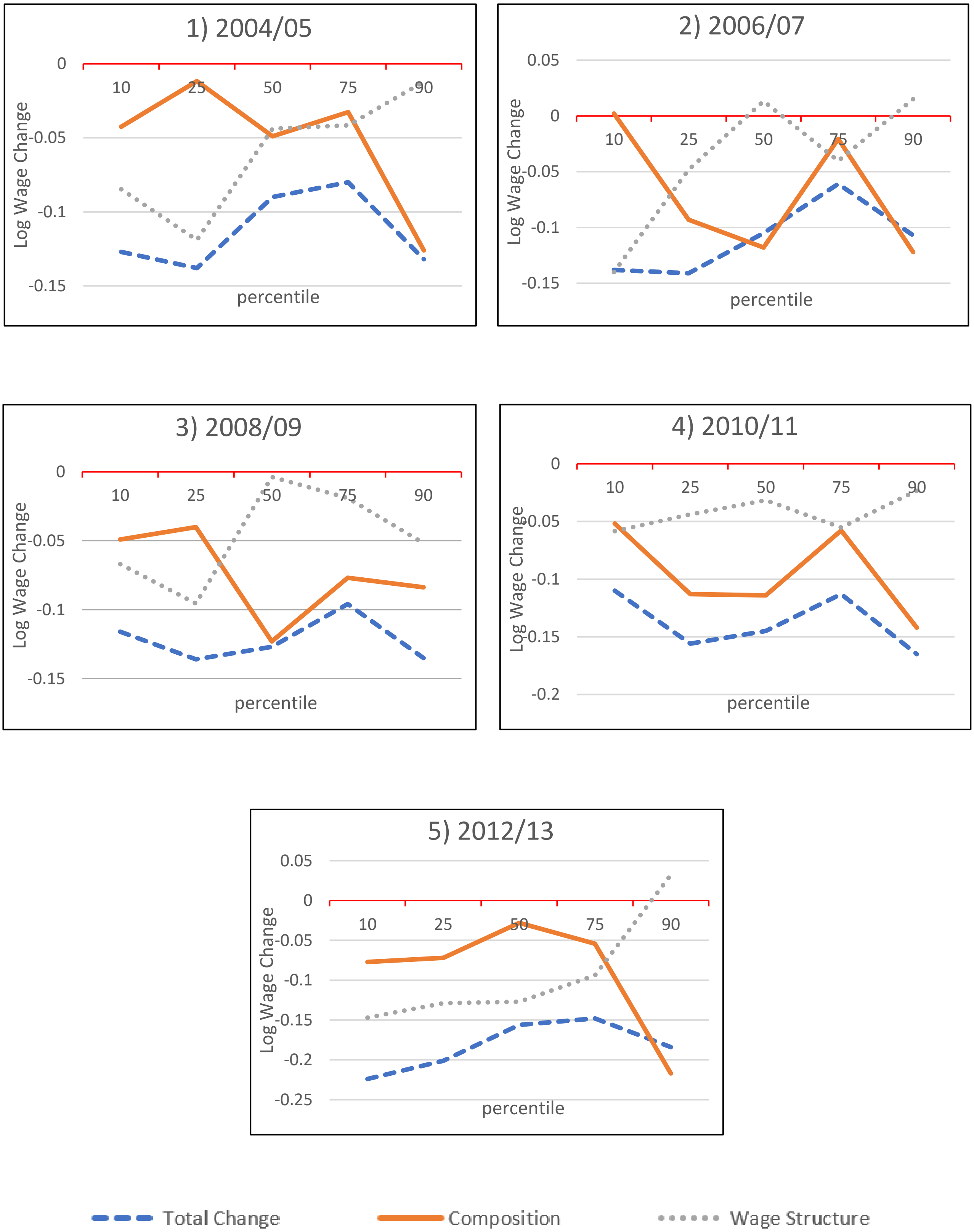


Figure 2.2.2

Overall decomposition effect for Move Stayer (Yearly)

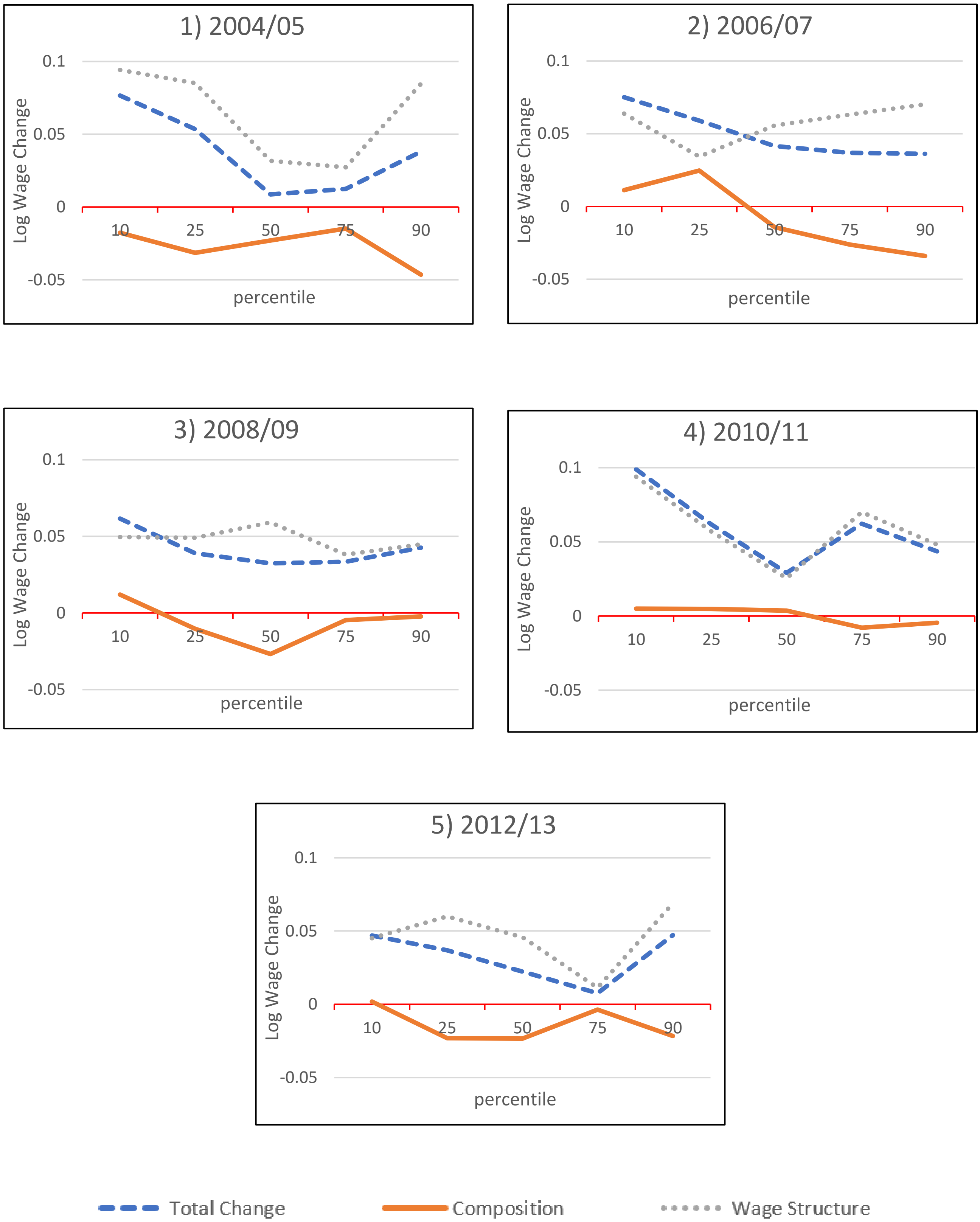


Figure 2.2.3

Overall decomposition effect for Non-Returning Double Mover (Yearly)

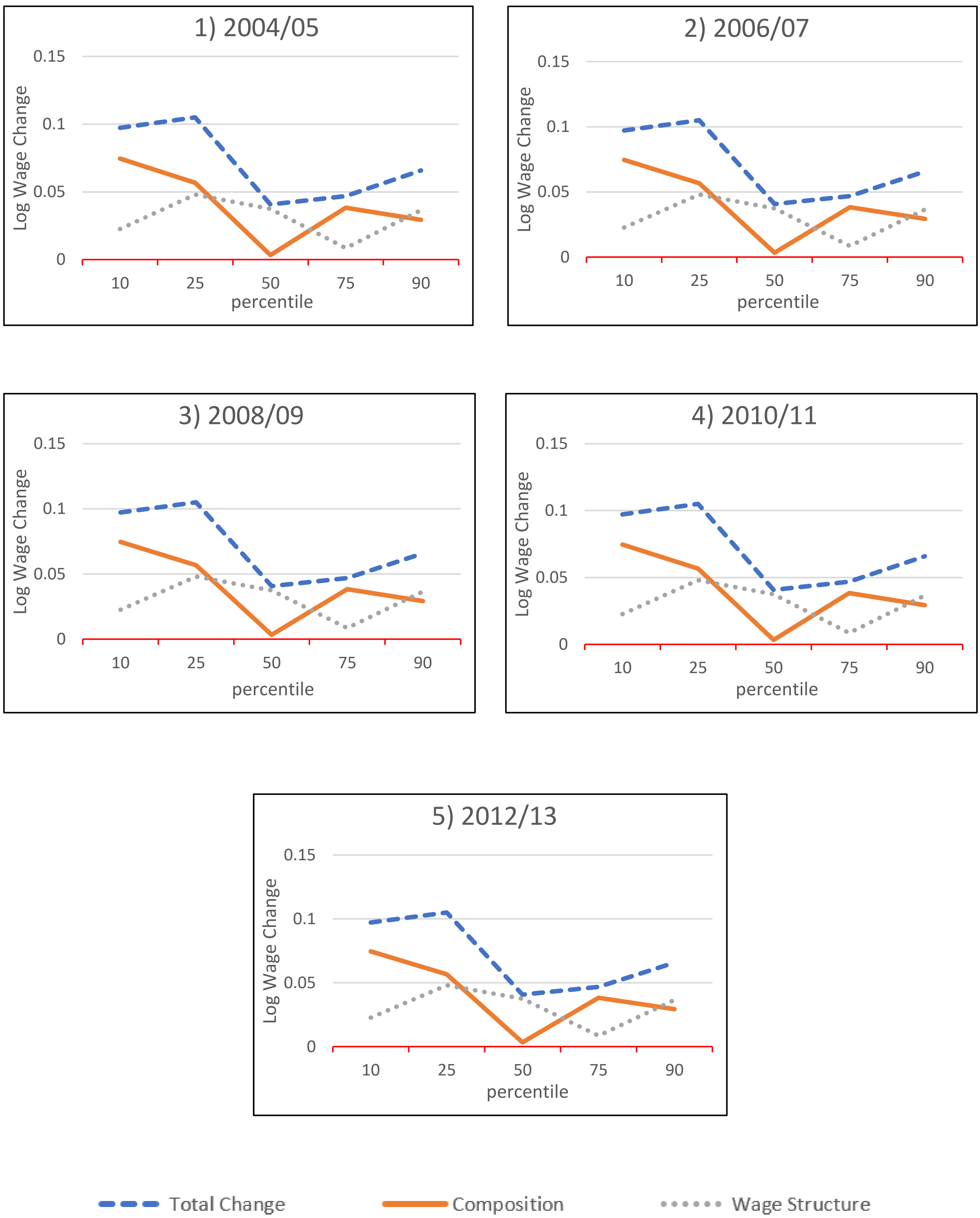


Figure 2.2.4

Overall decomposition effect for Stay Mover (Yearly)

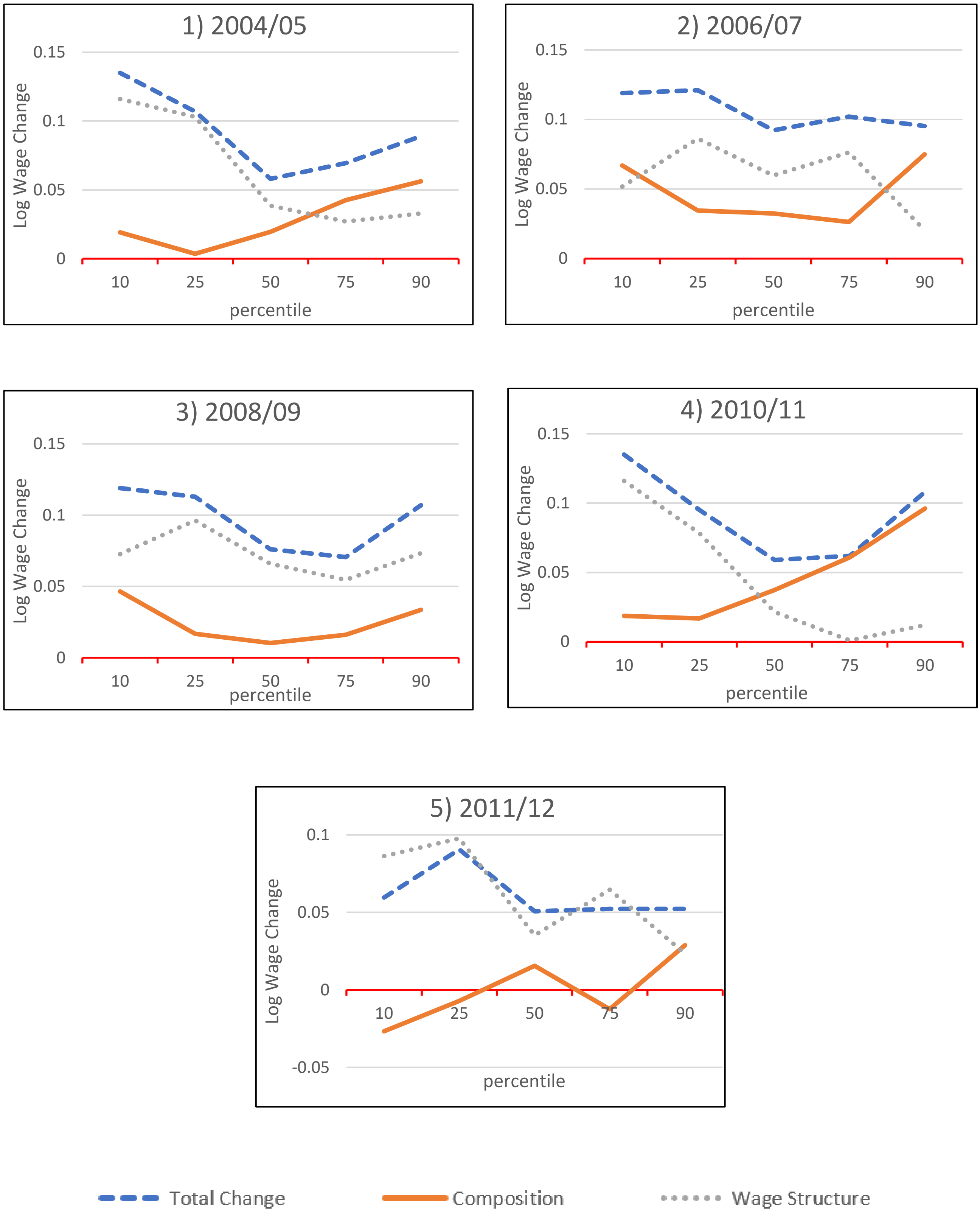


Table 2.5.1

Regression of aggregate decompositions on percentile for Move Returner

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Overall effect	-0.127***	-0.138***	-0.116***	-0.110***	-0.224***
Composition	-0.043	-0.002	-0.049	-0.052	-0.077
Wage Structure	-0.085*	-0.140	-0.067	-0.058	-0.147*
			25th Percentile		
Overall effect	-0.138***	-0.141***	-0.136***	-0.156***	-0.201***
Composition	-0.012	-0.093	-0.040	-0.113*	-0.072
Wage Structure	-0.119***	-0.048	-0.096*	-0.044	-0.129*
			50th Percentile		
Overall effect	-0.090**	-0.105***	-0.127***	-0.145***	-0.156***
Composition	-0.049	-0.118	-0.123**	-0.114**	-0.028
Wage Structure	-0.044*	0.013	-0.004	-0.032	-0.127*
			75th Percentile		
Overall effect	-0.080*	-0.061***	-0.096***	-0.113***	-0.148***
Composition	-0.033	-0.021	-0.077**	-0.058	-0.054
Wage Structure	-0.041	-0.040	-0.019	-0.055	-0.094*
			90th Percentile		
Overall effect	-0.132*	-0.107***	-0.135***	-0.165***	-0.184***
Composition	-0.126	-0.122	-0.084	-0.142*	-0.217**
Wage Structure	-0.012	0.015	-0.052	-0.023	0.032
Observations	**3,984**	**4,588**	**7,095**	**9,481**	**4,899**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to two decimal places.

Table 2.5.2

Regression of aggregate decompositions on percentile for Move Stayer

Years	**2004/05**	**2006/07**	**2008/09**	**2010/11**	**2012/13**
			10th Percentile		
Overall effect	0.077***	0.075***	0.062***	0.099***	0.047*
Composition	-0.018	0.011	0.012	0.005	0.002
Wage Structure	0.094***	0.064*	0.050*	0.094***	0.045
			25th Percentile		
Overall effect	0.054***	0.059***	0.039***	0.062***	0.037*
Composition	-0.032	0.025	-0.010	0.005	-0.023
Wage Structure	0.085***	0.034	0.049**	0.057***	0.060*
			50th Percentile		
Overall effect	0.009	0.042***	0.032***	0.029***	0.022
Composition	-0.023	-0.014	-0.027*	0.004	-0.023
Wage Structure	0.032	0.056***	0.059***	0.026*	0.046*
			75th Percentile		
Overall effect	0.013	0.037**	0.034***	0.062***	0.008
Composition	-0.015	-0.026	-0.005	-0.008	-0.004
Wage Structure	0.027	0.063**	0.038*	0.070***	0.011
			90th Percentile		
Overall effect	0.038	0.036	0.043**	0.044**	0.047*
Composition	-0.047	-0.034	-0.002	-0.004	-0.022
Wage Structure	0.085*	0.070*	0.045	0.048*	0.069
Observations	**5,118**	**5,966**	**8,795**	**11,309**	**4,457**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to two decimal places.

Table 2.5.3

Regression of aggregate decompositions on percentile for (NRDM)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Overall effect	0.097***	0.104***	0.117***	0.131***	0.078***
Composition	0.075	0.071*	0.085***	0.035	0.111**
Wage Structure	0.023	0.033	0.032	0.097***	-0.034
			25th Percentile		
Overall effect	0.105***	0.108***	0.105***	0.114***	0.092***
Composition	0.057	0.071**	0.042*	0.015	0.081**
Wage Structure	0.048	0.037	0.063**	0.099***	0.011
			50th Percentile		
Overall effect	0.041***	0.077***	0.074***	0.055***	0.054***
Composition	0.003	0.031	0.021	0.009	0.044*
Wage Structure	0.037	0.046*	0.053***	0.046***	0.011
			75th Percentile		
Overall effect	0.047***	0.088***	0.074***	0.060***	0.060***
Composition	0.038	0.028	-0.002	0.055***	0.042
Wage Structure	0.009	0.060**	0.076***	0.005	0.018
			90th Percentile		
Overall effect	0.066***	0.073***	0.078***	0.102***	0.092***
Composition	0.029	0.028	-0.005	0.072**	0.006
Wage Structure	0.037	0.046	0.084**	0.030	0.086
Observations	**10,943**	**13,177**	**19,461**	**25,325**	**12,289**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.5.4

Regression of aggregate decompositions on percentile for Stay Mover

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Overall effect	0.135***	0.119***	0.119***	0.135***	0.060*
Composition	0.019	0.067**	0.047*	0.019	-0.027
Wage Structure	0.116**	0.052	0.073**	0.116***	0.086
			25th Percentile		
Overall effect	0.107***	0.121***	0.113***	0.095***	0.091***
Composition	0.004	0.035	0.017	0.017	-0.007
Wage Structure	0.103**	0.086***	0.096***	0.078***	0.098***
			50th Percentile		
Overall effect	0.058***	0.092***	0.076***	0.059***	0.051***
Composition	0.020	0.032*	0.010	0.037***	0.016
Wage Structure	0.038	0.060**	0.066***	0.022	0.035
			75th Percentile		
Overall effect	0.069***	0.102***	0.071***	0.062***	0.052***
Composition	0.043	0.026	0.016	0.061***	-0.012
Wage Structure	0.027	0.076***	0.055***	0.001	0.065*
			90th Percentile		
Overall effect	0.089***	0.095***	0.107***	0.108***	0.052*
Composition	0.056	0.075**	0.034	0.096***	0.029
Wage Structure	0.033	0.020	0.073**	0.012	0.023
Observations	**4,368**	**5,064**	**7,843**	**10,417**	**4,316**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Firstly, from Figure 2.1 (1), it is evident that during 2004/05, Move-Returner stands out as an anomaly. Move-Returner is the only mobility group that experiences a negative pay gap throughout the pay distribution. This implies that in general, graduates who are Move-Returner earn lower wages than Non-Mover across the pay distribution. Conversely, the other mobility groups display a positive wage premium throughout 2004/05, as shown on Figure 2.1 (2), 2.1 (3), and 2.1 (4). Except for Move-Returner, the composition effect and the wage structure effect are in general positive throughout the pay distribution. The only exception is the negative coefficients for the composition effect for Move-Stayer, although the results are statistically insignificant as shown in Table 2.5.2. Therefore, except for Move-Returner, graduates that migrate tend to earn higher wages compared to their non-moving counterparts.

Figure 2.2.1 to Figure 2.2.4 shows the breakdown of the overall effect for each mobility groups. As illustrated, the constituent of the overall effect is generated with the aggregation of the composition effect and the wage structure effect. For example, Figure 2.2.2 (1) shows that Move-Returner in 2004/05 both exhibit negative composition effect and wage structure effect. Particularly, it shows that the wage structure effect is considerably more negative at the lower end of the pay distribution compared to the composition effect. Table 2.5.1 indicates that at the 25th percentile, the wage structure effect is -11.9%, while the composition effect is -1.2%, resulting in a substantial gap of 10.7%. Conversely, at the 90th percentile, the wage structure effect is -1.2%, and the composition is -13.2%, resulting in a difference of 12%. Thus, on average, the wage structure (composition) effect predominantly influences the overall effect at the lower (higher) end of the pay distribution.

Moreover, it is noticeable that during 2004/05, all the mobility groups exhibit an overall negative correlation across the pay distribution. The largest overall pay gap decline observed is Stay-Mover, followed by Move-Stayer, Non-Returning Double Mover, and finally Move-Returner. The graphs shows that the composition effect line is primarily below the wage structure effect line. This implies that relatively, the composition effect drives down the overall effect of the wage compared to the wage structure effect.

2.5.1.4. Analysis of the subsequent years (2006/07 – 2012/13)

This section examines the overall decomposition for the years between 2006/07 – 2012/13 and compares to the 2004/05 base year. Firstly, graduates that are Move-Returner consistently maintain a negative pay gap across all percentile intervals for each subsequent year, as illustrated on Figure 2.1 (2) – Figure 2.1 (5) Furthermore, consistent with the base year, in general the composition effect and wage structure effect for Move-Returner are negative. For the few coefficients that are positive for Move-Returner (including overall, composition, and wage structure effect) they are all statistically insignificant as revealed on Table 2.5.1.

In contrast, throughout the pay distribution, no other migration group exhibits a negative wage premium. This reinforces the earlier observation that graduates migrating for higher education institution (HEI) and/or employment generally achieve higher earnings compared to Non-Mover graduates. Thus, as previously stated, Move-Returner is an anomaly as it is the only group which consistently exhibits a negative wage premium. This finding is consistent with previous research such as Faggian et al. (2007).

Moreover, compared to the wage structure effect, the coefficient magnitude of the composition effect is relatively low. This can be observed from Figure 2.2.2 – Figure 2.2.4 as the composition effect is relatively flat and coefficient is close to zero throughout the pay distribution. This is supported by Table 2.5.1 to Table 2.5.4 since the majority of the coefficients for the composition effect are not statistically significant. An explanation is that the differences in the covariates compared to the base group i.e. Non-Mover, are not sufficient. Only the composition effect with a large coefficient is likely to be statistically significant. For instance, many of the composition effects are significant for Move-Returner as the composition effect coefficient considerably deviates compared to Non-Mover.

Fourthly, it is evident from Figure 2.2.1 the composition effect for the subsequent years remains negatively correlated throughout the pay distribution. Interestingly, the decline in the composition effect intensifies over the years. For example, across the pay distribution, the composition effect decreased by 8.3% (-4.3 % to -12.6%) in 2004/05, whereas the composition effect decreased by 14% (-7.7% to -21.7%) in 2012/13. This indicates that as the percentile increases, the composition effect becomes more negative for Move-Returner graduates and this effect exacerbates as the years increases.

In contrast, the other three mobility groups (excluding Mover-Returner) maintain a positive wage premium across every percentile interval which is consistent to the trend of the base year. Notably, in the first four periods, Stay-Mover outperform Non-Returning Double-Mover due to a highly positive wage structure effect. However, this is reversed in 2012/13 as the Non-Returning Double-Mover wage premium exceeds Stay-Mover wage premium across the entire pay distribution. This is primarily due to a decline in the composition effect for Stay-Mover in the last period e.g. at the 75th percentile, the composition effect decreases from 6.1% in 2010/11 to -1.2% in 2012/13.

As an overall, the trend pattern for the three mobility groups closely resembles the base year results. Generally, the wage premium across the pay distribution shows a negative correlation for the three groups. This is anticipated since a significant number of graduates resides and works in London. Therefore, since London is known as a region of high earners, the London Non-Mover graduates may skew the earnings distribution. This effect is magnified at the top of the earnings percentile, thus a possible explanation for the negative correlation in wage premium for the three mobility groups.

Figure 2.2.1 to 2.2.4 demonstrates that, in general, the wage structure effect outweighs the composition effect across the mobility groups. This is supported by Table 2.5.1 to Table 2.5.4, as the coefficient is greater as an absolute value for the wage structure effect compared to the composition effect coefficient. Also, there are more wage structure effect results that are statistically significant compared to the composition effect, this is especially the case when the wage premium is relatively high. This finding is consistent with the results for the base year. The higher degree of statistical significance observed in the wage structure effect is expected, as the primary determinant for this analysis is the selection of graduate mobility rather than the characteristics of the individuals.

Lastly, for the mobility groups that experience a positive wage premium, the highest point of the overall change is usually at the 90th percentile as seen in Figure 2.1. Table 2.5.1 to Table 2.5.4 confirms that the results are nearly always highest at the 90th percentile for all the groups except for Move-Returner. This suggests the widest wage premium gap between the graduates that move and those that do not, occurs at the top 10% of the earning distribution spectrum.

2.5.1.5. Aggregate decomposition (Comparing between percentiles)

An alternative approach is by comparing the aggregate decomposition results across the percentile intervals by examining Figure 2.3 and Figure 2.4.1 to Figure 2.4.4.

The positions of each mobility group on Figure 2.3 are similar to those on Figure 2.1. Move-Returner consistently show negative figures, performing the poorest, followed by Move-Stayer. Non-Returning Double Mover and Stay-Mover both demonstrate better performance with relatively similar results. Notably, at each percentile interval, the line representing Stay-Mover on Figure 2.3 surpasses that of Non-Returning Double Mover. This indicates in the early periods, regardless of a graduate's position in the pay distribution, Stay-Mover consistently outperform Non-Returning Double Mover. For instance, in 2004/05, Stay-Mover’s wage premium exceeds Non-Returning Double Mover in every pay distribution. Over subsequent periods, the pay gap advantage for Stay-Mover persists compared to Non-Returning Double Mover, albeit diminishing in absolute value over time as discussed in the previous section.

Conversely, the positions of Non-Returning Double Mover and Stay-Mover reverses in the later periods across all pay distribution intervals. This shift is evident in Figure 2.3 and Table 2.5.3, where the pay gap of Non-Returning Double Mover exceeds Stay-Mover in 2012/13. This indicates that Stay-Mover outperform Non-Returning Double Mover in the early and intermediate periods of the study. However, in the latter period, Non-Returning Double Mover’s pay gap surpasses Stay-Mover, regardless of the position in the pay distribution.

Figure 2.3

Overall effect for each mobility type (Percentile)

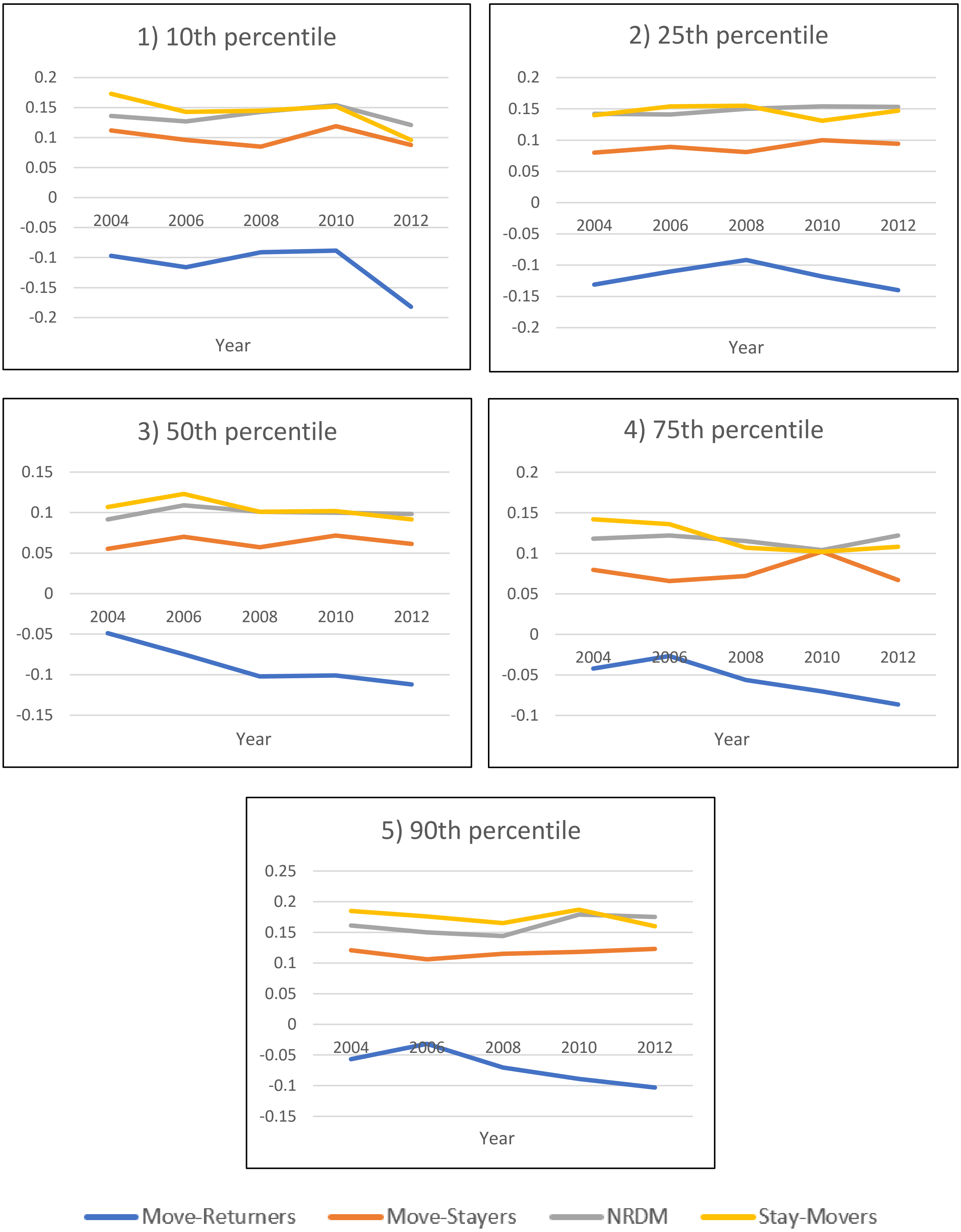


Figure 2.4.1

Overall decomposition effect for Move Returner (Percentile)

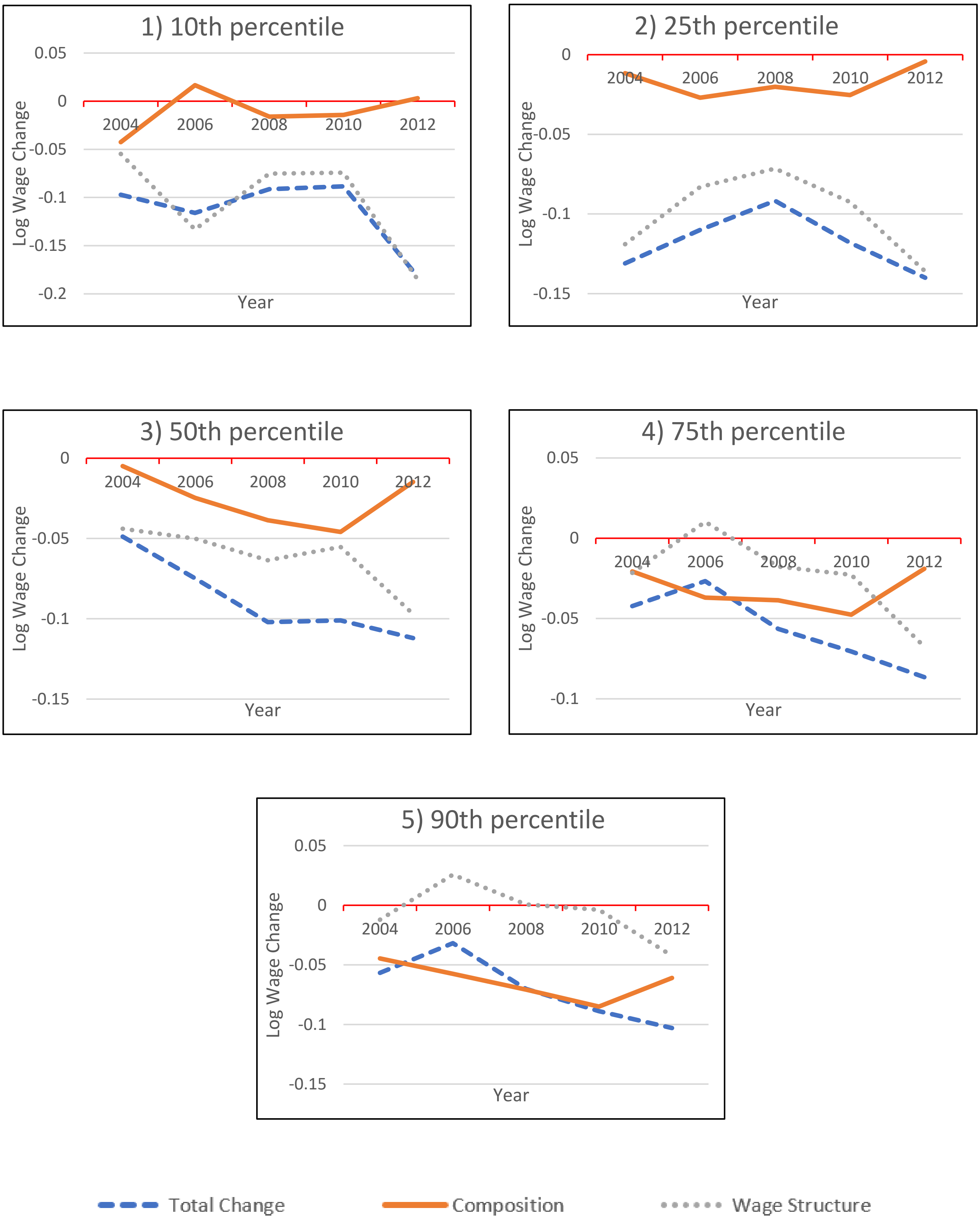


Figure 2.4.2

Overall decomposition effect for Move Stayer (Percentile)

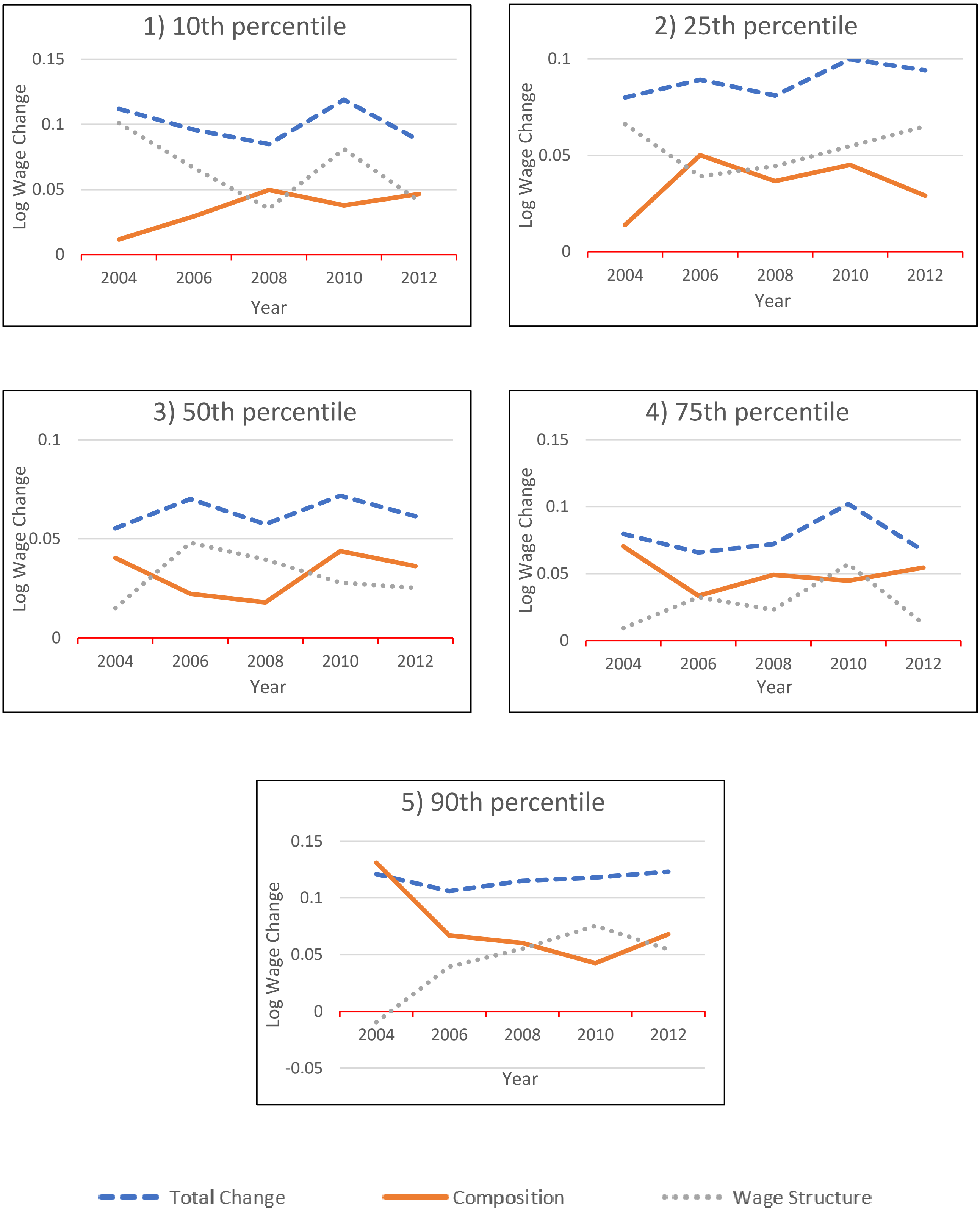


Figure 2.4.3

Overall decomposition effect for Non-Returning Double Mover (Percentile)

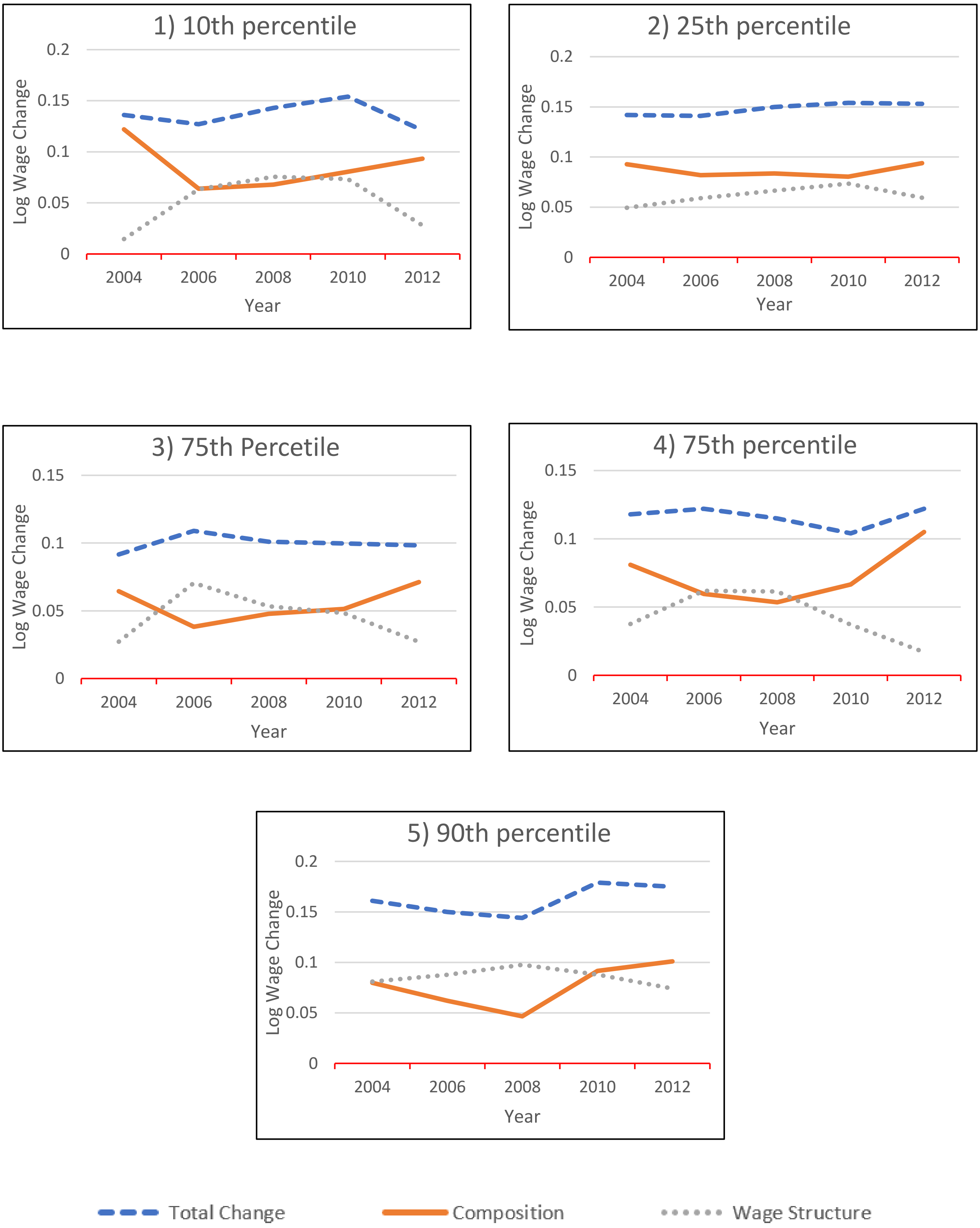
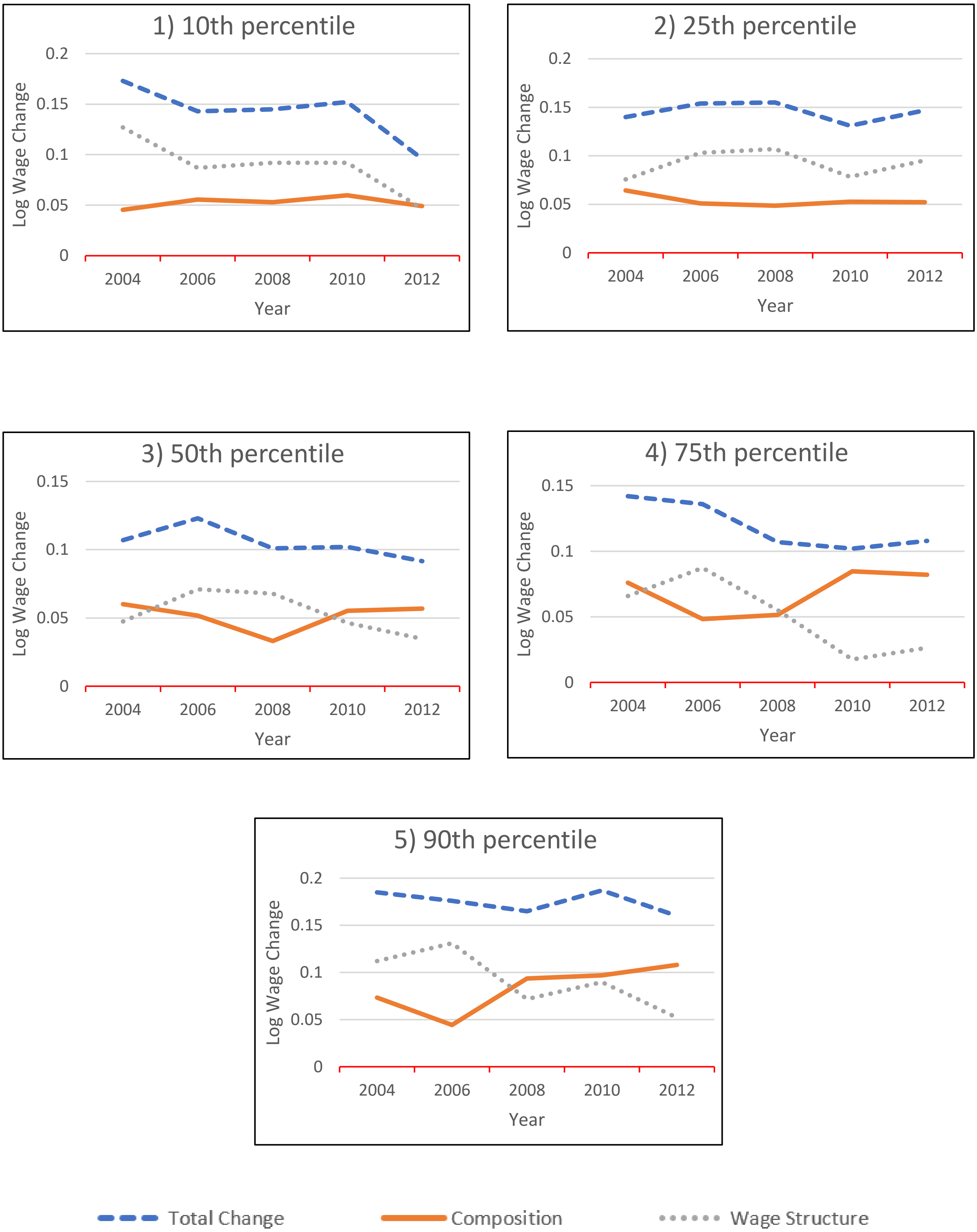


Figure 2.4.4

Overall decomposition effect for Stay Mover (Percentile)



For future research, it would be valuable to explore whether Non-Returning Double Mover consistently maintain a higher pay gap compared to Stay-Mover, or if the phenomenon observed in 2012/13 is a unique occurrence. Prior studies suggest that individuals with higher mobility often experience higher earning premium compared to their less mobile counterparts (Nigel O'Leary, 2017) due to the benefits of wider job search. This suggests that Non-Returning Double Mover should outperform Stay-Mover. However, Kidd, O'Leary and Sloane (2017) found that earning premiums predominantly arise from secondary migration, i.e., changes in employment.

Both Figure 2.3 (1) and Figure 2.4.1 (1) to Figure 2.4.1 (5) demonstrate that Move-Returner consistently show a downward trend with the wage premium as the percentile intervals increases. This indicates that the wage premium of Move-Returner's has progressively worsen as year increases. Particularly, there is a significant wage premium decline as the years increased which suggests that the disparity in earning premiums compared to Non-Mover is widening. For example, at the 10th percentile, the wage premium remained relatively indifferent between 2004/05 and 2010/11. However, 2012/13 experienced a substantial wage premium decrease by 11.5% compared to the previous year. Therefore, this finding indicates that the income inequality between Move-Returner and Non-Mover has worsen over time which contradicts the findings by Kidd, O'Leary, and Sloane (2017).

2.5.2. Decomposition based on RIF estimates

2.5.2.1. Aggregate composition effects

The aggregate composition effect refers to the change in the dependent variable (income) due to the changes in the composition of the mobility groups rather than the actual changes in income. The pure composition effect refers to the change in the average income solely due to changes in the composition of the population (i.e. migration group). Therefore, it isolates the effect of changes in the composition from any other factors influencing income. The specification error represents the misspecified when estimating the regression model or the set of independent variables. This can occur when the model fails to capture the true relationship between the dependent variable and the independent variables as a result of incorrect assumptions or omitting necessary variables.

Figure 2.5

Composition effect for each mobility type (Yearly)

Figure 2.6.1

Aggregate composition effect for Move Returner (Yearly)

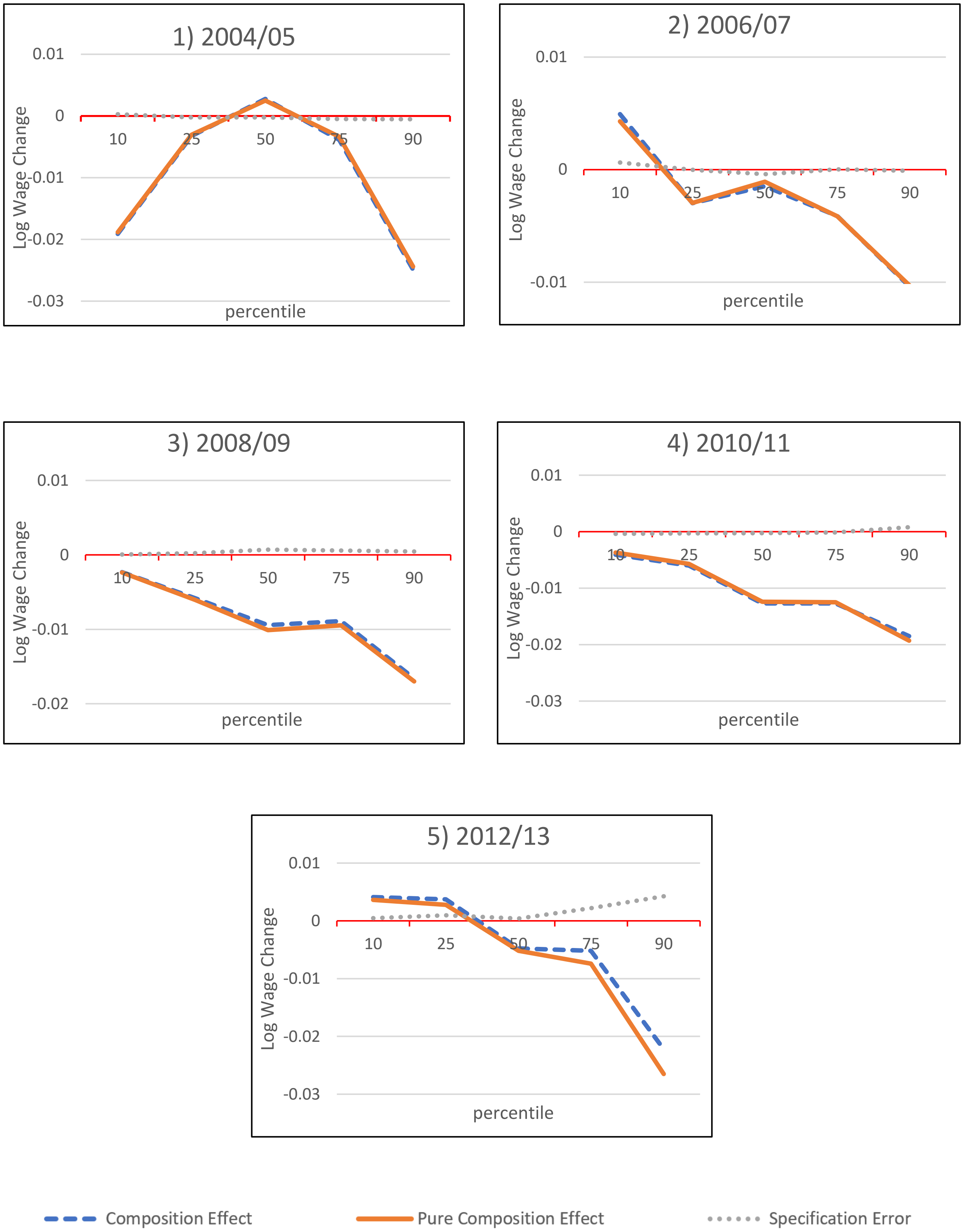


Figure 2.6.2

Aggregate composition effect for Move Stayer (Yearly)

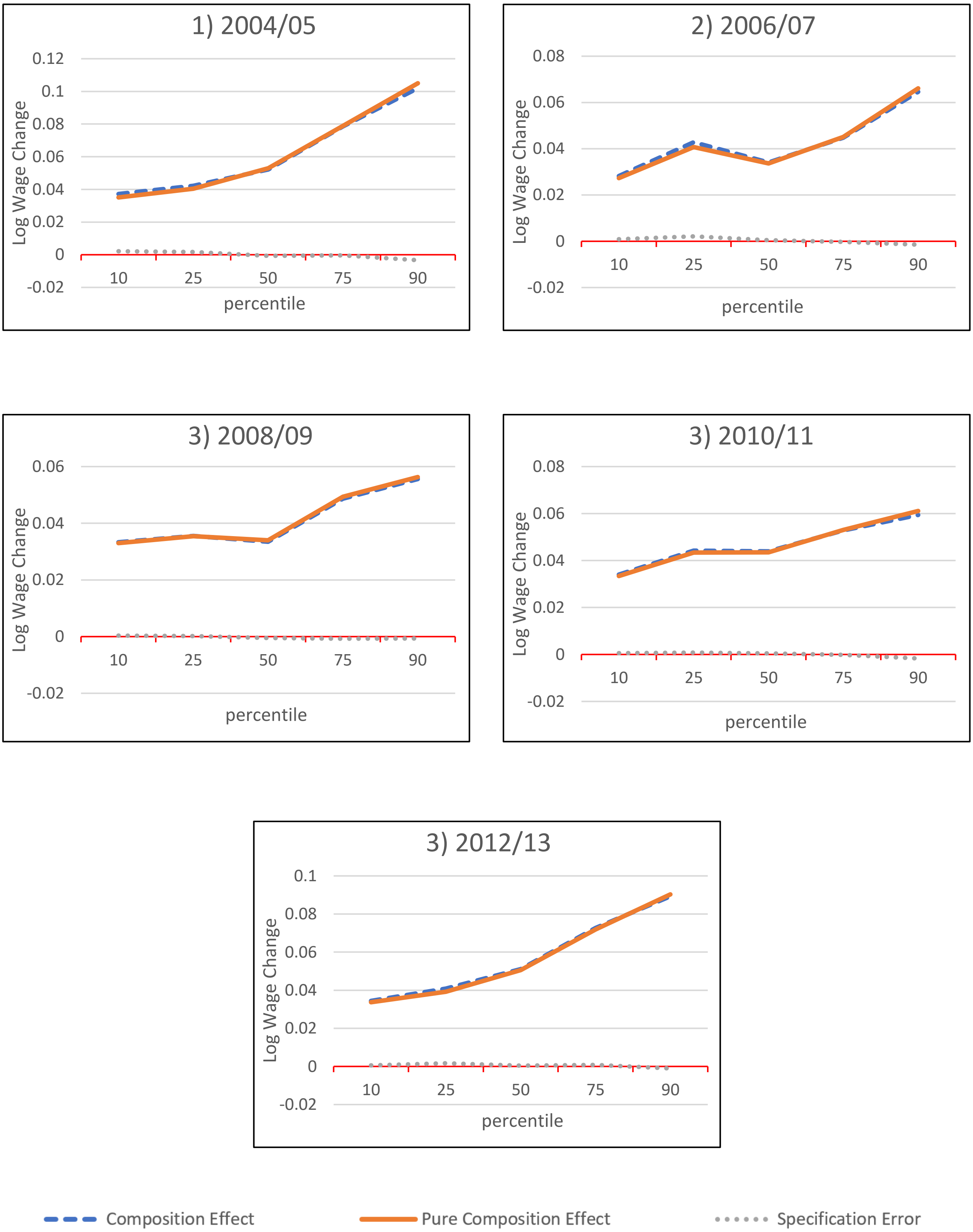


Figure 2.6.3

Aggregate composition effect for Non-Returning Double Movers (Yearly)

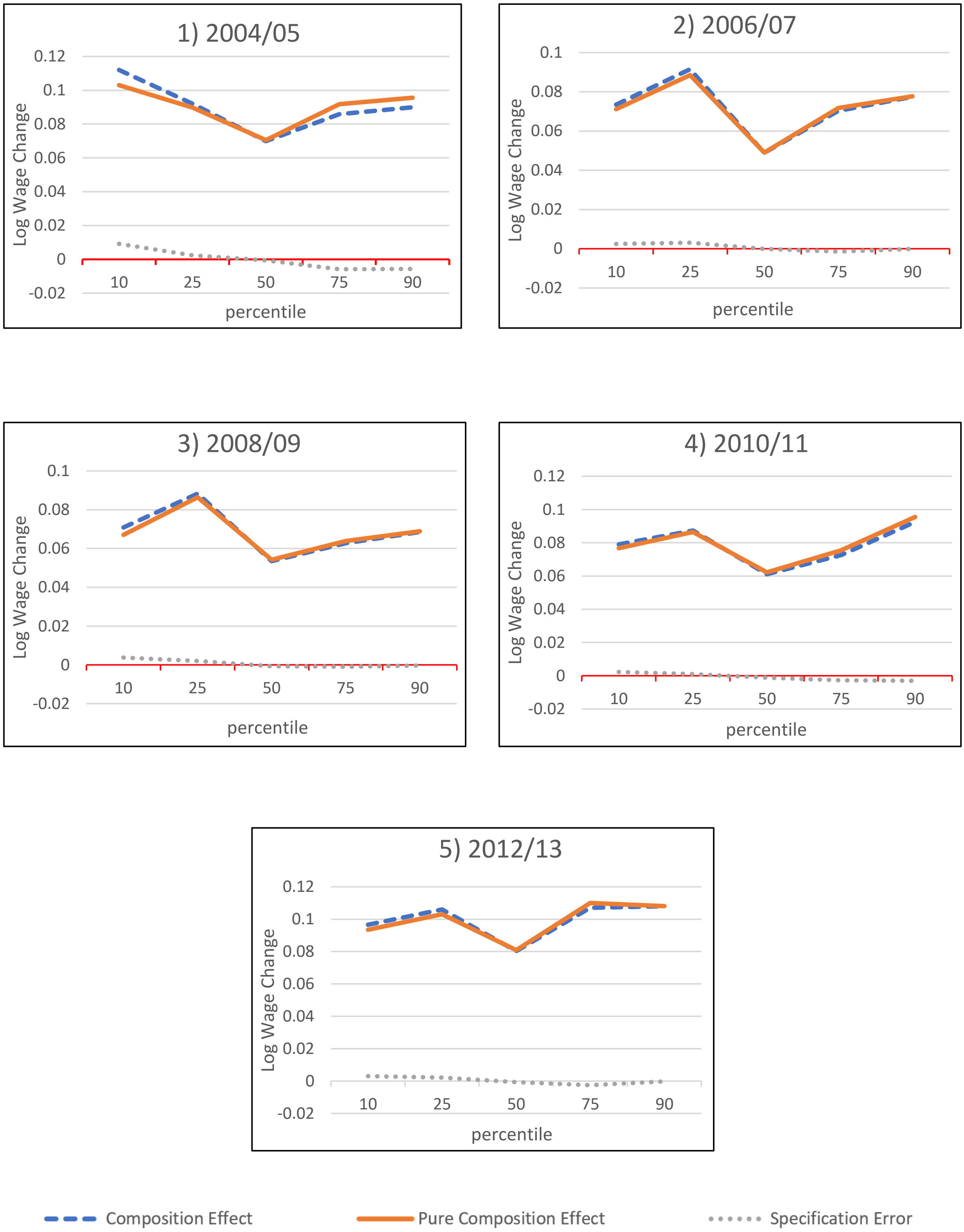


Figure 2.6.4

Aggregate composition effect for Stay Mover (Yearly)

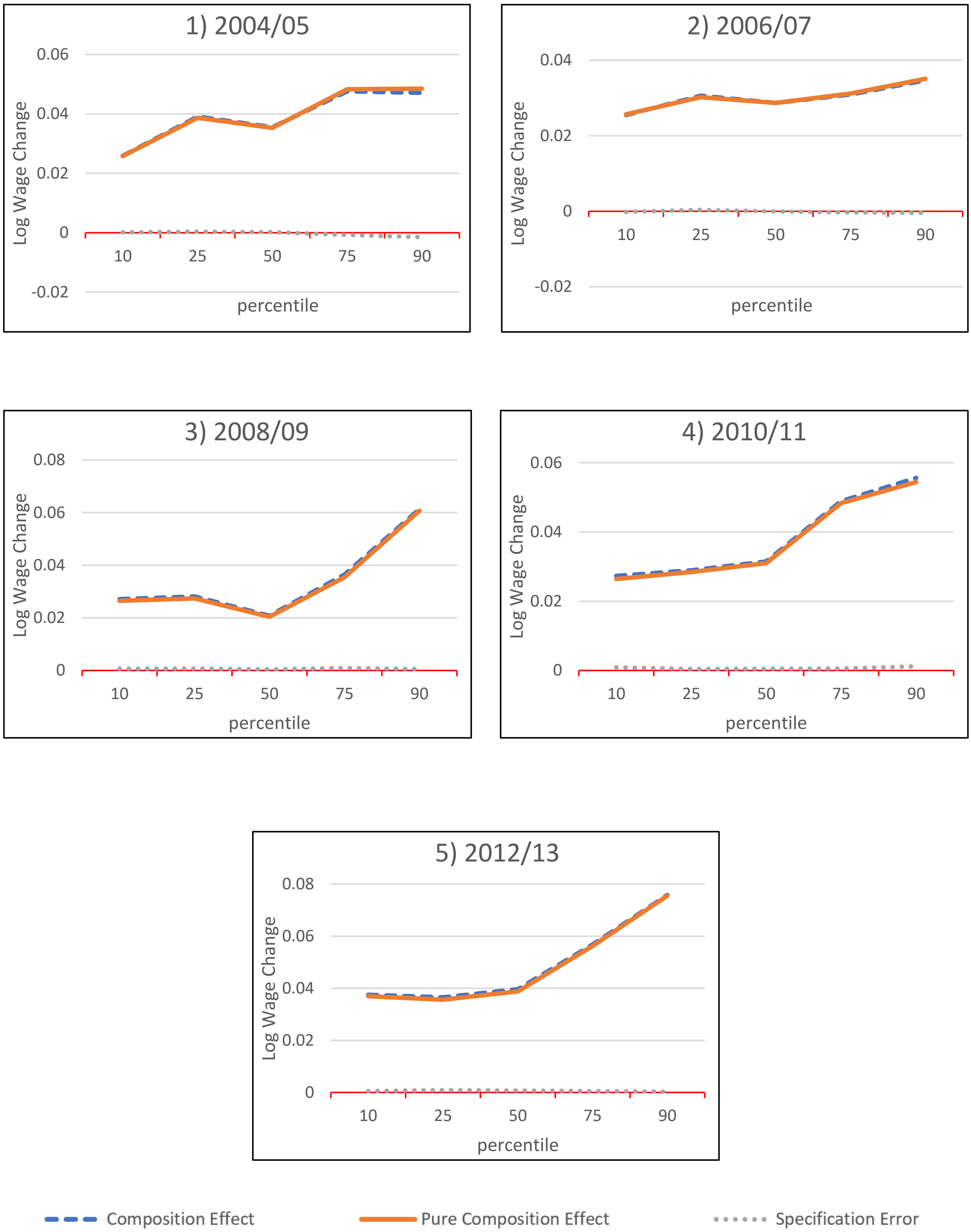


Table 2.6.1

Regression of composition effect on percentile for Move Returner

Years	2004/05	2006/07	2008/09	2010/11	2012/13
		10th Percentile			
Total Comp effect	-0.019	0.005	-0.003	-0.004	0.002
Pure Comp effect	-0.019***	0.004	-0.003	-0.004	0.002
Specification Error	0.000	0.001	-0.000	-0.000	0.000
		25th Percentile			
Total Comp effect	-0.003	-0.004	-0.006	-0.007	0.003
Pure Comp effect	-0.003***	-0.004	-0.006	-0.006	0.002
Specification Error	-0.000	0.000	-0.000	-0.000	0.001
		50th Percentile			
Total Comp effect	0.003	-0.002	-0.011	-0.013	-0.005
Pure Comp effect	0.003***	-0.002	-0.011	-0.013	-0.005
Specification Error	-0.000	0.000	0.000	-0.000	0.000
		75th Percentile			
Total Comp effect	-0.004	-0.004	-0.010	-0.013	-0.006
Pure Comp effect	-0.003***	-0.004	-0.010	-0.013*	-0.008
Specification Error	-0.000	-0.000	-0.000	-0.000	0.002
		90th Percentile			
Total Comp effect	-0.025	-0.011	-0.017	-0.019	-0.026
Pure Comp effect	-0.024***	-0.011	-0.017	-0.020*	-0.030**
Specification Error	-0.001	-0.000	0.000	0.001	0.004
Observations	**22,138**	**1,622**	**2,758**	**3,378**	**3,944**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.6.2

Regression of composition effect on percentile for Move Stayer

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Total Comp effect	0.031	0.025	0.029	0.0288	0.027
Pure Comp effect	0.029***	0.024**	0.029***	0.0288***	0.027**
Specification Error	0.001	0.001	0.000	0.000	-0.000
			25th Percentile		
Total Comp effect	0.033*	0.039**	0.031*	0.037**	0.032
Pure Comp effect	0.032***	0.037***	0.031***	0.036***	0.030***
Specification Error	0.001	0.002	-0.000	0.001	0.001
			50th Percentile		
Total Comp effect	0.039**	0.029**	0.029**	0.036***	0.039**
Pure Comp effect	0.039***	0.028***	0.029***	0.036***	0.039***
Specification Error	0.000	0.000	-0.000	0.000	0.000
			75th Percentile		
Total Comp effect	0.060***	0.036**	0.041***	0.043***	0.062***
Pure Comp effect	0.060***	0.036***	0.042***	0.043***	0.061***
Specification Error	-0.000	0.000	-0.000	-0.000	0.000
			90th Percentile		
Total Comp effect	0.071**	0.051*	0.045*	0.049**	0.072**
Pure Comp effect	0.073***	0.052***	0.044***	0.052***	0.073***
Specification Error	-0.002	-0.001	0.000	-0.003	-0.001
Observations	**3,528**	**4,378**	**6,158**	**7,034**	**3,060**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.6.3

Regression of composition effect on percentile for NRDM

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Total Comp effect	0.112***	0.073***	0.071***	0.076***	0.097***
Pure Comp effect	0.103***	0.071***	0.067***	0.074***	0.094***
Specification Error	0.009	0.002	0.004	0.002	0.003
			25th Percentile		
Total Comp effect	0.092***	0.090***	0.085***	0.083***	0.104***
Pure Comp effect	0.090***	0.087***	0.083***	0.082***	0.102***
Specification Error	0.002	0.003	0.002	0.001	0.002
			50th Percentile		
Total Comp effect	0.070***	0.048***	0.052***	0.058***	0.079***
Pure Comp effect	0.071***	0.048***	0.052***	0.059***	0.079***
Specification Error	-0.001	-0.000	-0.001	-0.001	-0.001
			75th Percentile		
Total Comp effect	0.086***	0.068***	0.059***	0.071***	0.104***
Pure Comp effect	0.092***	0.070***	0.060***	0.074***	0.106***
Specification Error	-0.006	-0.001	-0.001	-0.003	-0.002
			90th Percentile		
Total Comp effect	0.090***	0.075***	0.065***	0.091***	0.103***
Pure Comp effect	0.096***	0.075***	0.065***	0.093***	0.103***
Specification Error	-0.006	-0.000	-0.001	-0.003	0.000
Observations	**15,178**	**18,800**	**27,490**	**35,066**	**18,724**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.6.4

Regression of composition effect on percentile for Stay Mover

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Total Comp effect	0.023	0.026	0.026	0.024	0.031
Pure Comp effect	0.023*	0.026**	0.026***	0.023***	0.031*
Specification Error	0.000	0.000	-0.000	0.001	0.001
			25th Percentile		
Total Comp effect	0.031	0.030	0.024	0.025	0.031
Pure Comp effect	0.031**	0.030***	0.024***	0.024***	0.030***
Specification Error	0.000	0.000	0.000	0.001	0.001
			50th Percentile		
Total Comp effect	0.030*	0.026	0.017	0.029**	0.035*
Pure Comp effect	0.029***	0.026***	0.017***	0.028***	0.034***
Specification Error	0.001	-0.000	0.000	0.000	0.001
			75th Percentile		
Total Comp effect	0.041*	0.027	0.030*	0.044***	0.048**
Pure Comp effect	0.041***	0.027***	0.030***	0.044***	0.047***
Specification Error	0.000	-0.000	0.000	0.001	0.001
			90th Percentile		
Total Comp effect	0.040	0.035	0.051*	0.053**	0.062*
Pure Comp effect	0.041**	0.035**	0.050***	0.051***	0.061***
Specification Error	-0.000	-0.000	0.001	0.001	0.000
Observations	**2,028**	**2,574**	**4,254**	**5,250**	**2,778**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

From Figure 2.5, it is evident that Move-Returner composition effect performs the poorest since it is the only group which have negative coefficients. The Non-Returning Double-Mover performs the best, then the Move-Stayer and followed by the Stay-Mover. This is true throughout the pay distribution and for every period. Throughout each period, the despite Move-Stayer consistently possess higher composition effect than the Stay-Mover, the difference every year is negligible, especially at the lower section of the percentile as shown on Figure 2.5. However, there is a significant gap between the Non-Returning Double-Mover with the rest of the mobility group. Therefore, the composition effect/pure composition effect of the Non-Returning Double-Mover compared to its counterpart, have significantly higher income distribution due to the changes in the composition of the population, rather than the changes in individual incomes.

Figure 2.6.1 to Figure 2.6.4 shows that the specification error is exceptional small throughout the entire pay distribution. This is true for each of the mobility groups at every period. Since the specification error is very close to zero as reinforced on Table 2.6.1 to Table 2.6.4, the composition effect closely follows the pattern and trend of the pure composition effect. The specification error is important as it ensures whether the estimation provides an accurate approximation of the model. Therefore, as the specification error is very small and insignificant, the null hypothesis that the component is significantly different from zero can be rejected. This suggests that the RIF regression for the model provides a highly accurate estimate of the overall composition and wage structure in the study.

Furthermore, it is noticeable that the Move-Returner is the only migration group which experiences a negative composition effect. The composition effect for Move-Returner is negatively correlated with respect to an increase in percentile. This implies that as the pay distribution increases, the average income for Move-Returner decreases compared to the Non-Mover due to the changes in the composition of the population. Figure 2.6.1 (1) is an anomaly as it demonstrates a quadratic graph, whereas Figure 2.6.1 (2) to Figure 2.6.1 (5) shows a relatively straight negative correlation.

On a contrary, the other three mobility groups (exclude Move-Returner), exhibit positive composition effect. Also, the composition effect and the pure composition effect is positively correlated with respect to an increase in percentile. Therefore, the pure composition effect increase represents an increase in the income of the three mobility groups due to changes in the composition of the population, while holding the income levels of individuals constant. In other words, this shows that the three mobility groups have a greater impact on the observed income distribution than the changes in the actual income levels.

2.5.2.1. Aggregate wage structure effects

Figure 2.7 shows the wage structure effect/total wage effect for each mobility group. All four groups are shown in one graph in their respective years for ease of interpretation. It is apparent that the Move-Returner performs the poorest as it is the only group which experiences negative results (negative throughout the entire pay distribution and for all years). In the first four period, Stay-Mover have the highest wage structure effect as shown on Figure 2.7 (1) to Figure 2.7(4). However, in the last period (2012/13), Non-Returning Double-Mover wage structure effect overtakes Stay-Mover as shown on Figure 2.7 (5). Throughout the first four periods, Non-Returning Double-Mover wage structure effect slightly performs better than Stay-Mover, especially at the lower percentiles. Although, overall, the two groups wage structure effect is relatively similar.

From Figure 2.7, it is evident that in general, the reweighting error effect and pure wage effect are both negatively correlated. As a result, the total wage effect is negatively correlated with respect to percentile. This is true for all the mobility groups. The only exception is for Non-Returner as the total wage effect is mostly positively correlated. In fact, Move-Returner is the only group which exhibits a negative total wage effect. This is true throughout the entire pay distribution for every year and most of the coefficients for the total wage effect are significant at the 1% level as observed in Table 2.7.1 to Table 2.7.2.

Figure 2.8.1 to Figure 2.8.4 illustrates the decomposition of the wage structure effect which possess two constituents, the reweighting effect, and the pure wage effect. The wage structure effect refers to the impact of various factors (e.g. distance travelled) on wage differentials among different groups (type of migration), while controlling for differences in observable characteristics. While the reweighting error evaluates the quality of the reweighting strategy. The reweighting error should go to or be close to zero in large samples. A low reweighting error signifies a satisfactory quality of the reweighting. Finally, the pure wage structure refers to the portion of wage differentials that cannot be explained by observable characteristics e.g. education and age. In this research, this represents the unexplained variation in wages that is attributed by the migration by the graduates.

The total wage effect for Non-Returning Double Mover is below the pure wage effect and reweighting error since both the components are negative (total wage effect is the aggregation of both the components) as shown in Figure 2.8.1. Since the coefficients of the total wage effect is below zero for the Move-Returner group, a positive correlation would imply that as the percentile increases, the total wage effect gets closer to zero. For instance, during 2012/13, the total wage effect is closer to zero as the percentile increase. This suggests that at the upper section of the pay distribution, the total wage effect is closer to that of the Non-Mover group.

Moreover, it is noticeable from Figure 2.8.1 (1) – Figure 2.8.1 (5), at the lower section of the pay distribution, the wage structure effect is greater than the reweighting effect in terms of absolute value. This is reversed at the upper end of the pay distribution as the absolute value is greater for the pure wage effect compared to the reweighting error. This suggests at the lower end of the pay distribution, the pure wage effect is the predominate component that influences the total wage effect and vice versa at the upper end of the pay distribution. Also, the coefficients of the total wage effect in general increases in terms of absolute value as the year increases. For instance, the wage structure effect experiences the greatest magnitude in the negative direction during 2012/13 as shown on Table 2.7.1.

With the other three mobility groups, the trend pattern of the total wage effect highly follows the trend of the pure wage effect as shown on Figure 2.8.2 – Figure 2.8.4. This arises since the reweighting error is relatively flat and close to zero, although the error line is negatively correlated with respect to percentile. This suggests that as the percentile increases, the counterfactual is not as well identified compared to the lower percentile points. Although, as an overall, the reweighting error coefficients are not substantial, thus, will not negatively affect the wage structure effect results. Also, since the reweighting error is close to zero, the total wage effect is heavily influenced by the result of the pure wage effect.

Figure 2.7

Wage structure effect for each mobility type (Yearly)

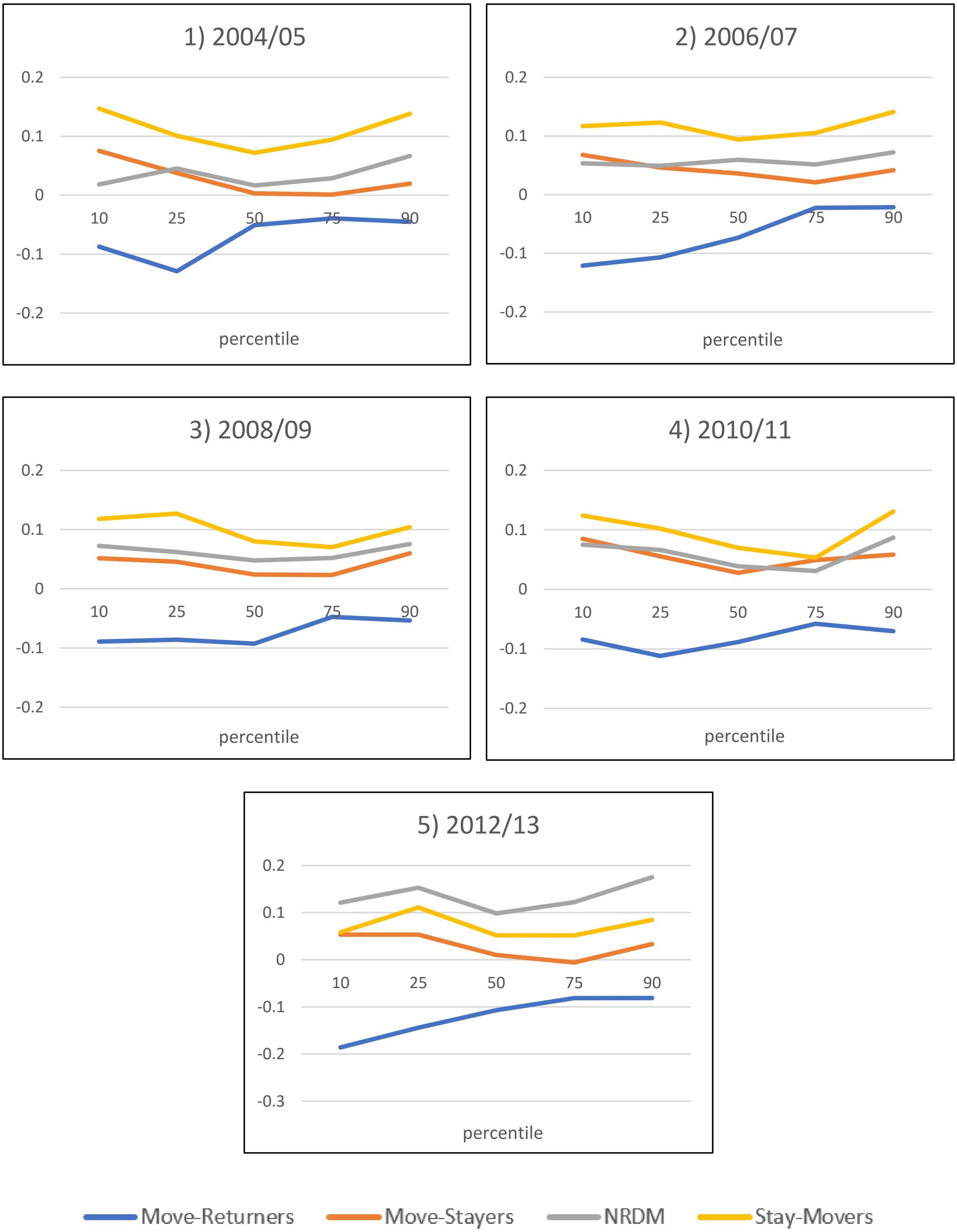


Figure 2.8.1

Aggregate wage structure effect for Move Returner (Yearly)

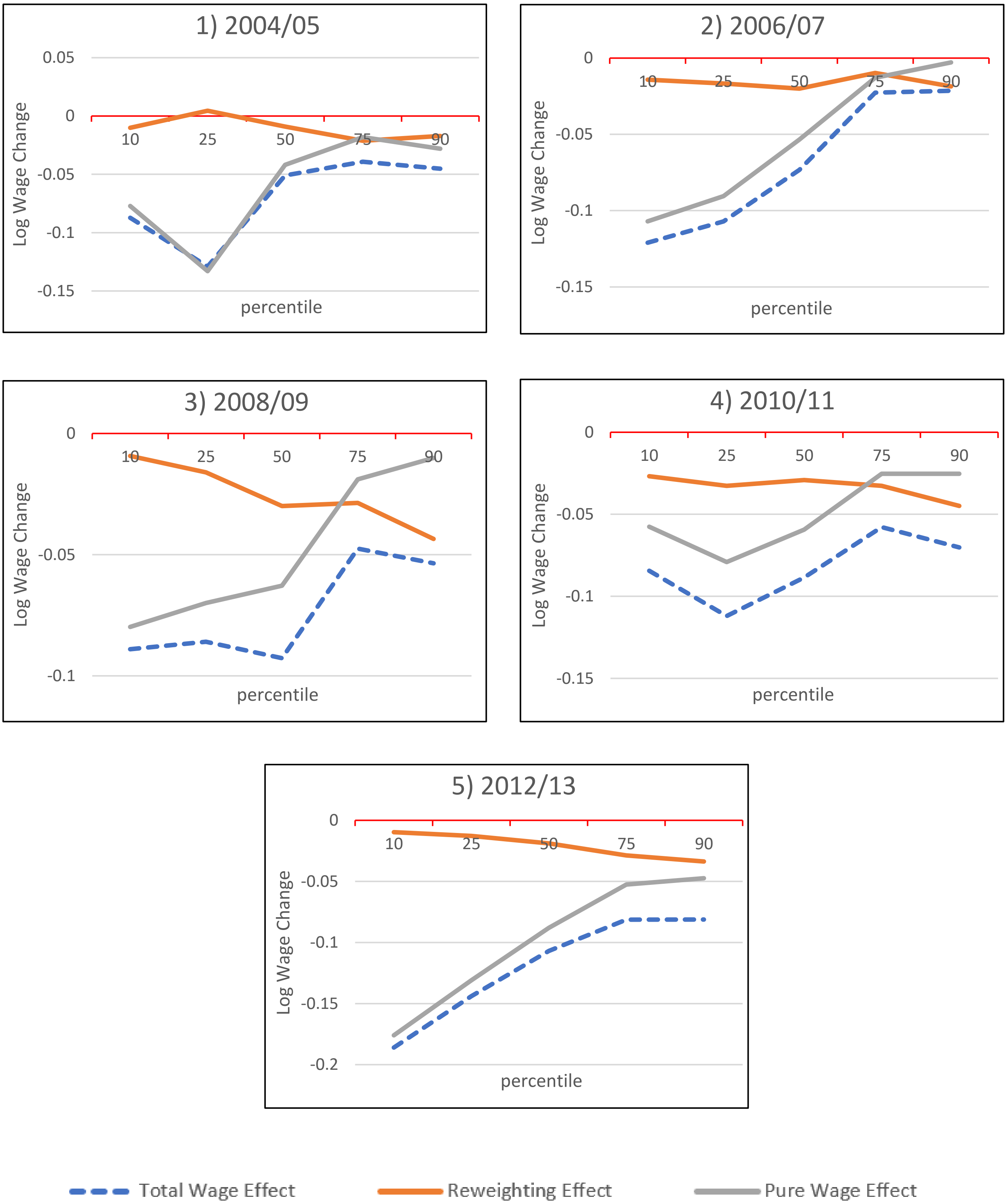


Figure 2.8.2

Aggregate wage structure effect for Move Stayer (Yearly)

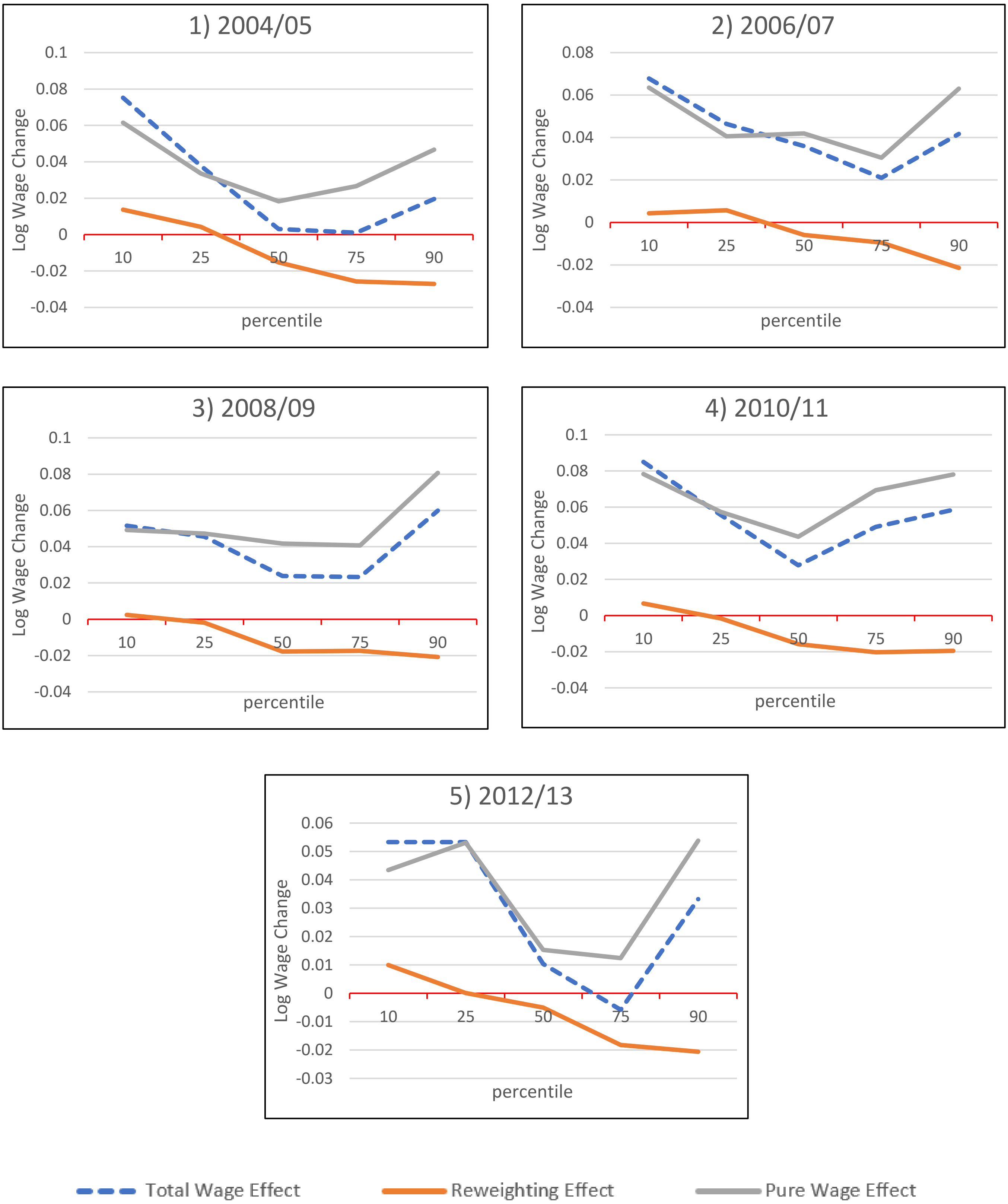


Figure 2.8.3

Aggregate wage structure effect for Non-Returning Double Movers (Yearly)

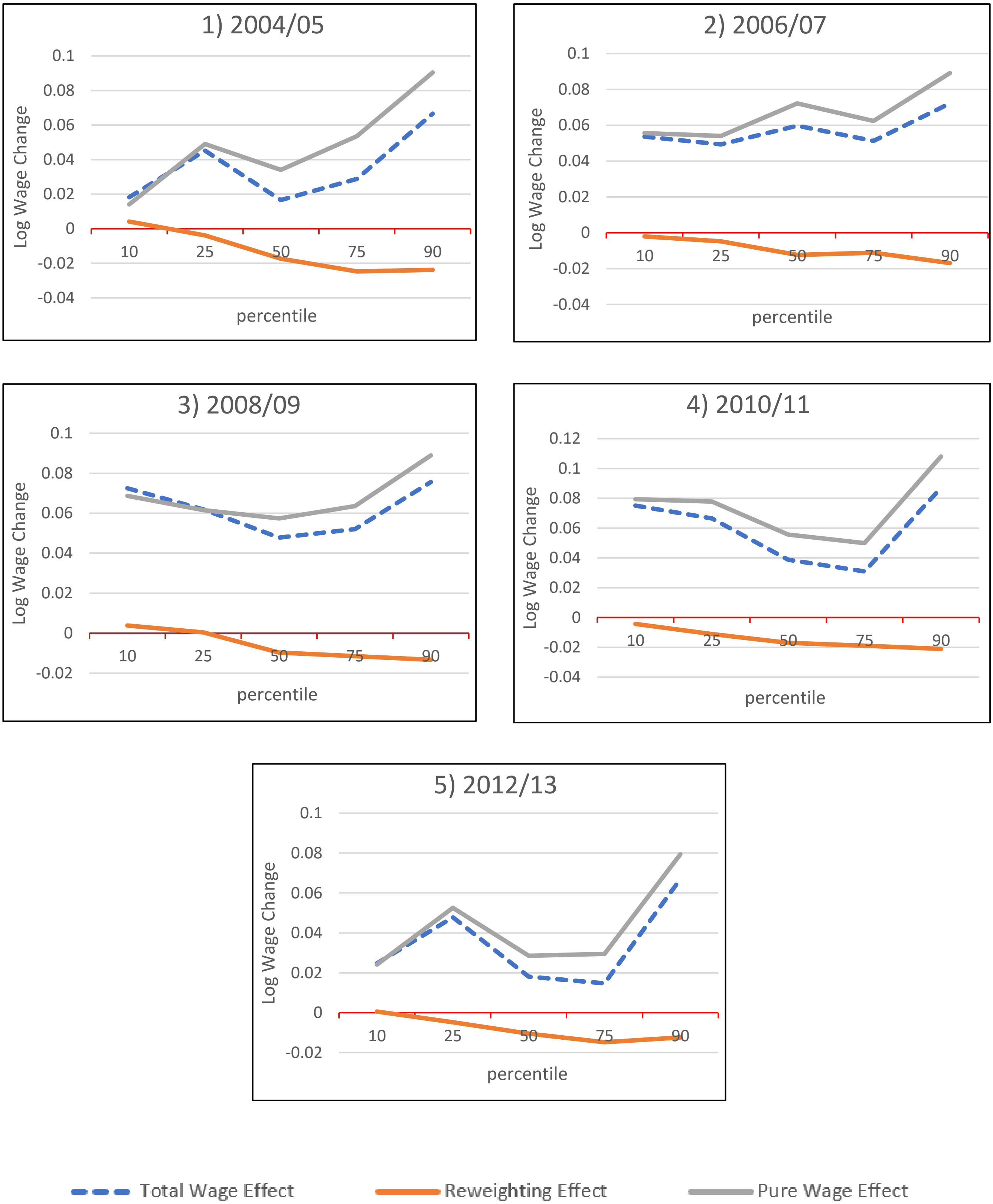


Figure 2.8.4

Aggregate wage structure effect for Stay Mover (Yearly)

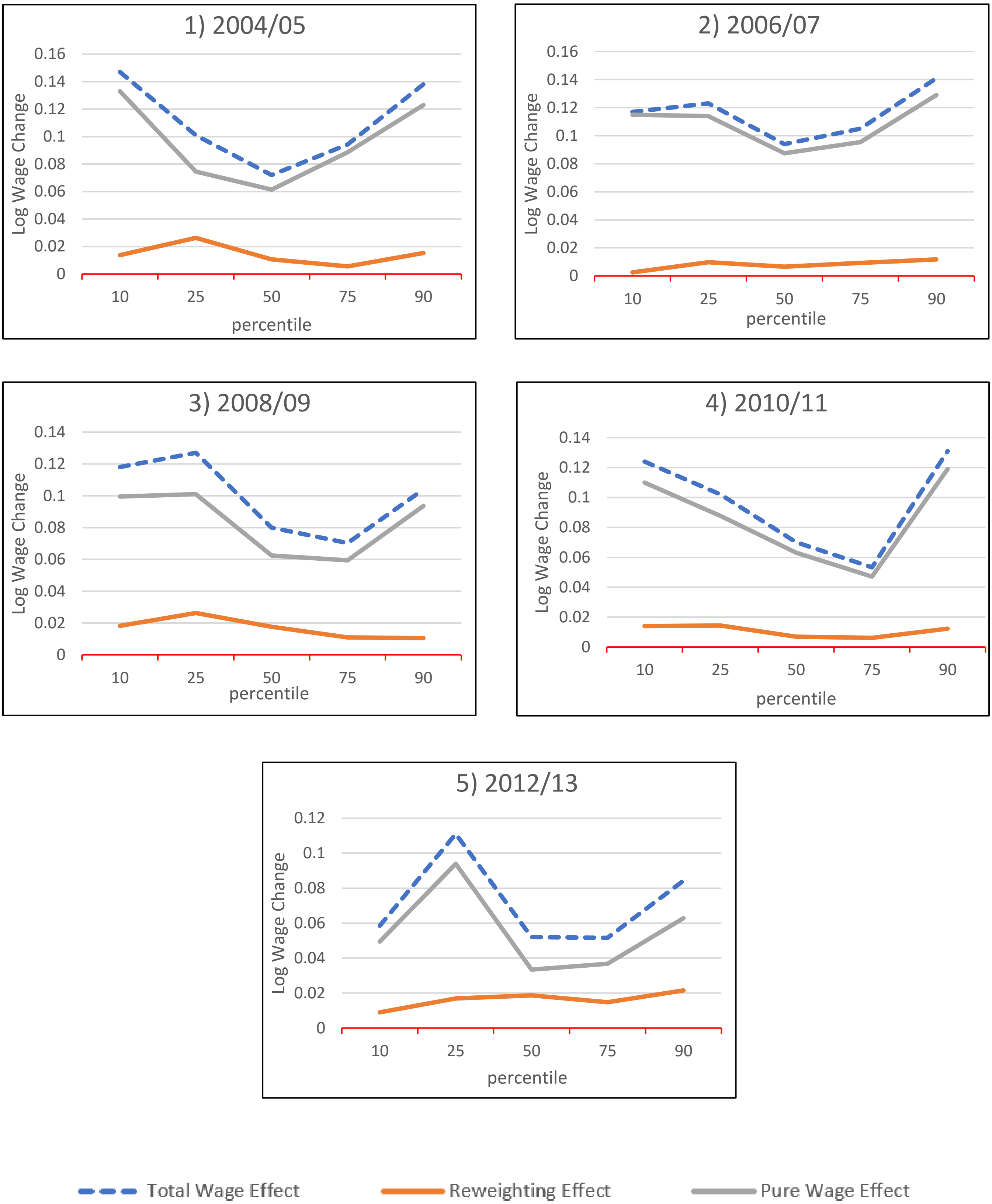


Table 2.7.1

Regression of wage structure effect on percentile for Move Returner

Years	**2004/05**	**2006/07**	**2008/09**	**2010/11**	**2012/13**
			10th Percentile		
Wage Structure effect	-0.108***	-0.143***	-0.113***	-0.106***	-0.226***
Reweighting Error	-0.017*	-0.031**	-0.032***	-0.046***	-0.046***
Pure Wage effect	-0.091	-0.112***	-0.081***	-0.060***	-0.181***
			25th Percentile		
Wage Structure effect	-0.134***	-0.138***	-0.129***	-0.150***	-0.204***
Reweighting Error	-0.037***	-0.051***	-0.060***	-0.073***	-0.064***
Pure Wage effect	-0.097***	-0.087***	-0.069***	-0.077***	-0.139***
			50th Percentile		
Wage Structure effect	-0.087**	-0.102***	-0.116***	-0.132***	-0.151***
Reweighting Error	-0.044***	-0.054***	-0.063***	-0.064***	-0.085***
Pure Wage effect	-0.043**	-0.048**	-0.053***	-0.068***	-0.066***
			75th Percentile		
Wage Structure effect	-0.076*	-0.057***	-0.086***	-0.100***	-0.142***
Reweighting Error	-0.054***	-0.047***	-0.078***	-0.075***	-0.094***
Pure Wage effect	-0.022**	-0.010	-0.009	-0.025*	-0.049***
			90th Percentile		
Wage Structure effect	-0.107	-0.096***	-0.119***	-0.146***	-0.158***
Reweighting Error	-0.061***	-0.081***	-0.099***	-0.100***	-0.096***
Pure Wage effect	-0.046**	-0.015	-0.020	-0.046*	-0.060**
Observations	**14,423**	**4,588**	**7,095**	**9,481**	**4,899**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.7.2

Regression of wage structure effect on percentile for Move Stayer

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.046**	0.051**	0.032*	0.070***	0.020
Reweighting Error	0.002	-0.006	-0.017**	-0.007	-0.023*
Pure Wage effect	0.044*	0.057**	0.050***	0.077***	0.043
			25th Percentile		
Wage Structure effect	0.020	0.020	0.008	0.025*	0.005
Reweighting Error	-0.030***	-0.019*	-0.037***	-0.032***	-0.035***
Pure Wage effect	0.050***	0.040**	0.045***	0.057***	0.040*
			50th Percentile		
Wage Structure effect	-0.030**	0.013	0.004	-0.007	-0.017
Reweighting Error	-0.045***	-0.033***	-0.047***	-0.041***	-0.056***
Pure Wage effect	0.015	0.046***	0.051***	0.034***	0.039**
			75th Percentile		
Wage Structure effect	-0.047***	0.001	-0.008	0.019*	-0.054***
Reweighting Error	-0.058***	-0.036***	-0.054***	-0.050***	-0.062***
Pure Wage effect	0.011	0.037**	0.046***	0.069***	0.008
			90th Percentile		
Wage Structure effect	-0.033	-0.015	-0.002	-0.005	-0.025
Reweighting Error	-0.050***	-0.063***	-0.063***	-0.052***	-0.059***
Pure Wage effect	0.017	0.048*	0.061***	0.047**	0.034
Observations	**5,118**	**5,966**	**8,795**	**11,309**	**4,457**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.7.3

Regression of wage structure effect on percentile for NRDM

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	-0.015	0.031*	0.046***	0.056***	-0.019
Reweighting Error	-0.020*	-0.018*	-0.019***	-0.021***	-0.034**
Pure Wage effect	0.006	0.048***	0.065***	0.077***	0.015
			25th Percentile		
Wage Structure effect	0.013	0.018	0.020*	0.031***	-0.013
Reweighting Error	-0.048***	-0.036***	-0.039***	-0.048***	-0.055***
Pure Wage effect	0.061***	0.054***	0.059***	0.079***	0.042***
			50th Percentile		
Wage Structure effect	-0.029***	0.029***	0.023***	0.003	-0.024**
Reweighting Error	-0.057***	-0.045***	-0.044***	-0.051***	-0.073***
Pure Wage effect	0.028***	0.073***	0.067***	0.048***	0.049***
			75th Percentile		
Wage Structure effect	-0.039***	0.020**	0.015*	-0.011*	-0.045***
Reweighting Error	-0.070***	-0.046***	-0.055***	-0.061***	-0.077***
Pure Wage effect	0.031**	0.066***	0.070***	0.050***	0.033***
			90th Percentile		
Wage Structure effect	-0.024	-0.002	0.014	0.011	-0.011
Reweighting Error	-0.076***	-0.075***	-0.067***	-0.074***	-0.078***
Pure Wage effect	0.052***	0.073***	0.081***	0.085***	0.067***
Observations	**10,943**	**13,177**	**19,461**	**25,325**	**12,289**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table 2.7.4

Regression of wage structure effect on percentile for Stay Mover

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.113***	0.093***	0.094***	0.111***	0.028
Reweighting Error	-0.001	-0.011	-0.002	0.000	-0.019
Pure Wage effect	0.114***	0.104***	0.095***	0.110***	0.047
			25th Percentile		
Wage Structure effect	0.076***	0.090***	0.089***	0.071***	0.060***
Reweighting Error	-0.020*	-0.017*	-0.005	-0.016**	-0.024**
Pure Wage effect	0.096***	0.107***	0.094***	0.087***	0.084***
			50th Percentile		
Wage Structure effect	0.028*	0.066***	0.059***	0.030***	0.016
Reweighting Error	-0.030***	-0.020**	-0.014**	-0.022***	-0.043***
Pure Wage effect	0.058***	0.086***	0.073***	0.053***	0.059***
			75th Percentile		
Wage Structure effect	0.029*	0.076***	0.040***	0.018	0.004
Reweighting Error	-0.031***	-0.020**	-0.025***	-0.029***	-0.049***
Pure Wage effect	0.060***	0.095***	0.065***	0.047***	0.053***
			90th Percentile		
Wage Structure effect	0.049*	0.061**	0.056**	0.056***	-0.009
Reweighting Error	-0.026*	-0.037**	-0.033***	-0.028***	-0.047***
Pure Wage effect	0.075**	0.097***	0.090***	0.084***	0.038
Observations	**4,368**	**5,064**	**7,843**	**10,417**	**4,316**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

The pure wage effect on average is positively the greatest for Stay-Mover, this in turn translates the group to possess the highest total wage effect as observed on Figure 2.8.4. Although, the positive gap of the total wage effect seems to diminish as the year increases. For instance, it is noticeable that the positive total wage effect (pure wage effect) is greatly reduced while comparing 2006/07 (Figure 2.8.4 (2)) and 2012/13 (Figure 2.8.4 (5)). In fact, this outcome of a decline in the total wage gap as the year increases is consistent with the other three groups (exclude Move-Returner). This suggests that as the year increases, the total wage effect (pure wage effect) gap reduces relative to the Non-Mover group.

In fact, during 2012/13, the Non-Returning Double-Mover experienced a negative total wage effect throughout the pay distribution as shown on Figure 2.8.3 (5).Interestingly, Move-Stayer experiences a positive total wage effect at the lower section of the pay distribution and transitions to negative at the higher section of the distribution. This is consistent in every year as shown in Figure 2.8.2. This is contributed by a relatively flat though positive pure wage effect coupled with a negatively correlated reweighting error. Often at the upper section of the percentile, the absolute value of the reweighting error is greater than the pure wage effect, thus, this results to the outcome of the total wage effect to transition from positive to negative as the pay distribution increases.

In summary, the results indicate as the pay distribution increase, in general, the total wage effect is less effected regardless of the mobility group of interest. As a result, the impact of whether a graduate is a Non-Mover, or a Mover is not so impactful if they are categorised at the top spectrum of earners. However, for individuals that are in the lower to the median bracket of the pay distribution, the type of graduate migration does make a significant difference to their total wage effect. The results show that on average, Move-Returner performs the poorest while Stay-Mover perform the best in terms of the pure wage effect and the total wage effect/wage structure effect.

2.6. Conclusion

2.6.1. Summary

This study examined trends in the earnings of young higher education (HE) graduates in the UK from 2004 to 2013, focusing on how migration decisions impact wage outcomes. The findings affirm that, with the exception of Move-Returners, there is a positive wage premium associated with migration to a preferred HE location and then to a region of employment upon graduation. This wage premium, derived through the Oaxaca-Blinder (OB) decomposition and recentred influence function (RIF) regression models, highlights the critical role of dual mobility decisions—both the initial migration for HE and subsequent employment migration—in shaping earning potential. These estimations reveal that both observed and unobserved characteristics of individual backgrounds, such as motivation, incentives, and aspirations, significantly influence migration choices and, consequently, wage outcomes in the labour market.

From this study, migration is positively associated with higher wages for most graduate categories, with one notable exception: the "move-returner" group, where the association is negative. This suggests that individuals who leave their original region for education or employment and later return may experience a wage penalty compared to those who either remain in their home region or migrate permanently to more economically robust areas. This finding raises important policy considerations, especially in the context of regional development and the UK's broader industrial strategy.

In line with prior research, the study found that migration generally results in a wage premium for graduates, with the Non-Returning Double-Mover and Stay-Mover groups experiencing the highest wage gains. Notably, the Stay-Mover group initially outperformed the Non-Returning Double-Mover group, but this advantage reversed in later periods, suggesting that the second stage of migration plays a decisive role in securing higher wages. This pattern is particularly relevant for those who extend their job search area by migrating, as it naturally broadens access to more and potentially higher-paying job opportunities. The analysis of Move-Returners and Move-Stayers confirms that the initial migration for HE alone does not confer a significant wage advantage; rather, it is the employment-related migration that strongly impacts wage premiums. Conversely, the Move-Returner group exhibits a negative wage premium, performing the poorest in 2012/13, with wages lower than those of Non-Movers. This negative premium likely stems from graduates returning to their domicile region for employment, which often offers limited job options and lower wages compared to regions with more robust job markets.

Further decomposition of wage determinants revealed that, with the exception of Move-Returners, the wage premium for other mobility groups is primarily driven by the wage structure effect rather than the composition effect, indicating that wage benefits are less about individual characteristics and more about the structural opportunities available in specific regions. The specification error of the composition effect and the reweighting error in the wage structure effect are minimal, validating the reliability of the RIF regression in accurately reflecting income distribution dynamics. These results underscore that both pure composition and wage structure effects largely drive the observed wage premiums, reflecting the desired robustness of this analytical approach.

2.6.2. Policy implications

The findings yield valuable policy implications, underscoring the need for a multifaceted approach to combat brain drain, bolster regional economic resilience, and promote equitable career opportunities across the UK. To address brain drain in economically weaker regions, targeted development initiatives could stimulate local job markets by attracting industries aligned with the skills of local graduates, making these areas more appealing for long-term careers. Investments in transport and digital infrastructure would further enhance competitiveness, supporting remote work options and reducing the impact of physical distance for graduates seeking quality employment. Additionally, higher education institutions (HEIs), working with policymakers, could tailor degree offerings to better match regional labour demands. By developing industry-specific training, especially in growing fields like renewable energy and digital services, and by expanding partnerships for apprenticeships and placements, HEIs can align education outcomes with local economic needs, improving job placement and encouraging graduates to remain where they studied.

The negative wage impact on the move-returner group suggests that policies encouraging migration should be balanced with measures that incentivise graduates to stay in or return to economically weaker regions without suffering wage penalties. A possible policy solution is to introduce targeted incentives to retain graduates in high-wage areas or encourage returnees to lower-wage regions. For example, wage subsidies or tax incentives for firms in less prosperous areas could make them more attractive to graduates. Alternatively, tailored policies could provide career advancement opportunities or relocation support for returnees, mitigating the wage penalties associated with returning to these regions.

This issue connects to a broader challenge within the UK's Industrial Strategy: balancing regional equality with the need to address brain drain. Brain drain, often seen as the outflow of skilled workers from economically weaker areas to metropolitan centres, can deepen regional disparities and hinder productivity growth in underperforming regions. While migration can help balance regional labour markets, the effects of brain drain indicate a need for policies that incentivise skilled workers to stay in or return to less advantaged regions. These policies must balance migration benefits with the need for regional development and equitable prosperity.

Given wage disparities across demographic groups, targeted support for underrepresented graduates is vital. Offering mentorship, career counselling, and networking opportunities to women, ethnic minorities, and graduates from low-income backgrounds can help address structural inequalities and improve access to well-paid jobs, especially for those facing higher mobility barriers. Government-backed initiatives to bridge these gaps would ensure that more graduates from diverse backgrounds benefit from migration-related wage premiums and regional job opportunities.

Policies encouraging graduates to return to their home regions could offer substantial benefits as well. Loan forgiveness or tax incentives for graduates who work in high-need areas could boost talent retention in underserved regions. Regional governments might also develop graduate retention programmes offering financial incentives—such as housing subsidies, relocation support, or professional development grants—for those choosing to stay or return. These programmes could target high-demand fields, ensuring graduates have access to fulfilling and competitive employment.

The findings on migration patterns also highlight the need to consider capital-labour complementarities in productivity discussions. Migration can redistribute labour to areas where it's most needed, potentially boosting productivity in underdeveloped regions. For these regions to fully benefit, however, complementary investments in physical and human capital are essential. These investments include infrastructure improvements, educational initiatives, and job creation in high-skill sectors aligned with the talents of migrating or returning graduates.

Empowering local governments to set long-term economic strategies could strengthen regional economies and build distinct economic identities. Developing industry clusters in sectors like tech, healthcare, or green energy would promote collaboration and innovation, attract graduates, and support sustainable local economies. Furthermore, investing in quality-of-life improvements—such as affordable housing, healthcare, and community amenities—would make these regions more appealing to young professionals who prioritise lifestyle alongside career growth.

Ultimately, these policy recommendations underscore the need for balanced efforts to tackle regional and demographic disparities in wage outcomes. By combining local economic incentives, targeted educational alignment with labour demands, support for underrepresented graduates, and initiatives to encourage graduate return migration, policymakers can create a more inclusive and dynamic national economy. This approach supports graduate aspirations while strengthening regional economies across the UK, fostering a sustainable framework for equitable economic growth.

In conclusion, policies managing graduate migration should address the labour market dynamics revealed by this study, considering both the regional wage penalties for returnees and the broader challenges of regional inequality and brain drain. Balancing migration incentives with strategies to reduce regional disparities and promote long-term productivity growth is essential for a comprehensive and effective industrial strategy.

2.6.3. Limitations

While this dataset offers a comprehensive scope for examining graduate migration and earnings outcomes, several limitations should be considered. The UK based dataset for this chapter accounts a total of 348,538 graduates. Following the removal of observations with missing data and the data cleaning process, a total of 133,472 observations remains, each containing complete migration histories. The DLHE survey, HESA asserts a response rate exceeding 75% (HESA, 2016). Assuming this claim is accurate, over 20% of graduates have not participated in the DLHE survey. Given the high proportion of non-participation, there's speculation that more successful graduates in securing desirable employment are more likely to respond. Conversely, less successful graduates may provide the postcode of a company's head office rather than their actual workplace location.

Due to a lack of sample size, particularly for the ethnic minority groups, certain factors influencing earnings such as industry-specific trends or regional variations, are not explicitly accounted for. For example, the investigation for regional effects is not feasible since the coefficient results lacked statistical power. Consequently, the analysis may not capture the full complexity of the determinants shaping wage differentials among graduate migrants.

Furthermore, using postcodes to measure migration proximity presents several challenges, primarily related to accuracy, which can lead to measurement errors. Additionally, it's important to note that the HESA survey does not collect information on the graduate's home address at the time of graduation; instead, it only provides the postcode of the student's permanent address prior to entering the program of study.

The postcodes of the three locations (i.e. domicile, higher education and employment) are recorded with only the first half of the postcode. Consequently, this diminishes the precision of location data, leading to measurement inaccuracies. For example, areas of domicile are broadly categorised, resulting in both affluent and disadvantaged regions being grouped together, thereby distorting the data. Similarly, this can happen with the location of study, as Higher Education Institutions (HEIs) are frequently clustered near one another, such as in Cardiff, London, and Swansea.

Another issue pertains to the method of recording migration information, particularly regarding the location of employment. The DLHE and LDLHE surveys capture employment locations approximately 6 months and 3.5 years after graduation, respectively. Consequently, any migration that occurs between these two points remains unobserved. For example, a graduate might relocate away from their place of study and domicile for employment and then return temporarily before settling elsewhere. This situation could lead to misclassification, such as labelling the graduate as a 'move-stayer' instead of a 'double non-returning mover.

Furthermore, graduates who pursued studies through long-distance learning programs often do not physically attend the university. Typically, distance education involves correspondence courses where students interact with the university through online platforms and postal communication. For instance, graduates of institutions like the Open University may be incorrectly categorized as 'Mover' in the second stage of migration, namely from university to employment.

Another issue arises with Higher Education Institutions (HEIs) that have multiple campuses, as the data may only reflect the main campus rather than the one where the graduate primarily studied. This discrepancy can lead to inaccuracies in tracking the locations of graduates during and after their studies. Although researchers may occasionally identify the correct geographic location of campuses, overall, this presents a significant limitation with potential measurement errors in the data.

Additionally, the survey fails to account for students who commute to university, often due to financial constraints. Many students, especially those studying in major urban centres like London, choose to reside outside the city centre to save costs and commute to their university. This commuting trend is particularly prevalent among part-time students or those balancing part-time work, resulting in discrepancies between the location of study and the student's place of residence. Consequently, this disparity complicates matters and increases the likelihood of measurement errors in postcode data. Hence, a graduate working in the city centre of London may not necessarily be categorized as a 'non-mover' in the second stage of migration.

Furthermore, the analysis focuses exclusively on graduates from their first degree in Higher Education Institutions (HEIs), omitting those with alternative educational backgrounds or pathways. This narrow scope may overlook important segments of the labour force, such as individuals with vocational qualifications or those pursuing further education beyond their initial degree. Consequently, the findings may not fully reflect the diversity of educational experiences and their corresponding impacts on earnings outcomes.

Moreover, the negative correlation observed in the total wage effect across all mobility groups, as depicted in Figure 2.7, underscores the complex relationship between migration and income outcomes. While the wage structure effect predominates at the lower end of the income distribution, the influence of migration diminishes at higher percentiles, suggesting that other factors may play a more significant role in determining earnings disparities among high-income earners.

2.6.4. Future studies

Extending the time frame beyond 2012 could provide valuable insights into the long-term trends and impacts of graduate mobility on income differentials. Analysing data up to the present day could reveal whether the observed patterns persist or evolve over time, especially in response to changing economic conditions and policy interventions[9].

Additionally, expanding the scope of analysis to include other datasets or sources of information could enrich our understanding of the factors driving income disparities among graduate migrants. Incorporating data from other national databases or international sources could offer comparative perspectives and shed light on the broader socio-economic context influencing earnings outcomes. Future research could delve deeper into the underlying mechanisms and drivers shaping the observed trends. Exploring the role of factors such as job location and industry composition.

Furthermore, conducting qualitative studies, such as interviews or surveys, with graduate migrants themselves could provide valuable insights into their decision-making processes,

[9] Investigating during and after the Covid-19 period would be insightful.

experiences, and perceptions regarding income disparities and mobility outcomes. Understanding the motivations, challenges, and aspirations of graduate migrants firsthand could complement quantitative analyses and offer a more comprehensive understanding of the phenomenon.

Lastly, exploring the potential implications of graduate migration patterns on broader socio-economic outcomes, such as regional development, labour market dynamics, and social mobility, could have significant policy implications. By examining the broader impact of graduate mobility beyond individual earnings, policymakers can develop more targeted and effective interventions to support inclusive growth and equitable distribution of opportunities.

2.7. Appendix

Table A.2.1.1

Regression of aggregate decompositions on percentile for Move Returner (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	-0.093***	-0.126***	-0.093***	-0.085***	-0.196***
Reweighting Error	-0.011	-0.013	-0.014	-0.033***	-0.017
Pure Wage effect	-0.082**	-0.113***	-0.079***	-0.052***	-0.179***
			25th Percentile		
Wage Structure effect	-0.130***	-0.115***	-0.098***	-0.113***	-0.147***
Reweighting Error	0.010	-0.026	-0.025*	-0.032***	-0.016*
Pure Wage effect	-0.139***	-0.089***	-0.073***	-0.081***	-0.131***
			50th Percentile		
Wage Structure effect	-0.059**	-0.075***	-0.091***	-0.092***	-0.118***
Reweighting Error	-0.016	-0.026*	-0.030***	-0.033***	-0.026**
Pure Wage effect	-0.043*	-0.049**	-0.061***	-0.059***	-0.092***
			75th Percentile		
Wage Structure effect	-0.046*	-0.028	-0.057***	-0.067***	-0.082***
Reweighting Error	-0.021*	-0.016	-0.035***	-0.038***	-0.030***
Pure Wage effect	-0.025	-0.012	-0.022	-0.029*	-0.052***
			90th Percentile		
Wage Structure effect	-0.050	-0.029	-0.057**	-0.075***	-0.088***
Reweighting Error	-0.024	-0.028	-0.046***	-0.041***	-0.039**
Pure Wage effect	-0.034	-0.001	-0.011	-0.034	-0.049*
Observations	**2,784**	**3,559**	**5,547**	**7,373**	**3,799**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.1.2

Regression of aggregate decompositions on percentile for Move Stayer (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.087***	0.077***	0.054***	0.098***	0.052*
Reweighting Error	0.0144	0.008	0.004	0.017	0.011
Pure Wage effect	0.073**	0.069***	0.050**	0.085***	0.041
			25th Percentile		
Wage Structure effect	0.041*	0.057**	0.050***	0.062***	0.057***
Reweighting Error	0.002	0.015	-0.007	-0.003	0.006
Pure Wage effect	0.039*	0.042**	0.057***	0.067***	0.051**
			50th Percentile		
Wage Structure effect	0.009	0.041**	0.026*	0.039***	0.013
Reweighting Error	-0.028*	-0.017	-0.029***	-0.028***	-0.017
Pure Wage effect	0.037	0.058***	0.055***	0.047***	0.030
			75th Percentile		
Wage Structure effect	0.002	0.027	0.027*	0.053***	-0.015
Reweighting Error	-0.033***	-0.014	-0.026***	-0.020***	-0.026*
Pure Wage effect	0.035*	0.041**	0.053***	0.071***	0.011
			90th Percentile		
Wage Structure effect	0.026	0.049*	0.062***	0.067***	0.037
Reweighting Error	-0.032**	-0.027*	-0.027**	-0.028*	-0.028
Pure Wage effect	0.058*	0.076***	0.089***	0.089***	0.065*
Observations	**3,783**	**4,895**	**7,189**	**9,104**	**3,298**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.1.3

Regression of aggregate decompositions on percentile for NRDM (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.026	0.054***	0.078***	0.086***	0.036
Reweighting Error	0.005	-0.005	0.007	-0.005	0.003
Pure Wage effect	0.021	0.059***	0.071***	0.091***	0.033
			25th Percentile		
Wage Structure effect	0.057***	0.053***	0.066***	0.071***	0.059***
Reweighting Error	-0.009	-0.015	0.005	-0.012**	-0.018
Pure Wage effect	0.066***	0.068***	0.061***	0.083***	0.077***
			50th Percentile		
Wage Structure effect	0.028	0.062***	0.059***	0.048***	0.021
Reweighting Error	-0.022***	-0.014**	-0.017**	-0.026***	-0.015*
Pure Wage effect	0.050***	0.076***	0.076***	0.064***	0.036**
			75th Percentile		
Wage Structure effect	0.037**	0.057***	0.053***	0.037***	0.028
Reweighting Error	-0.038***	-0.016**	-0.017***	-0.024***	-0.02**
Pure Wage effect	0.075***	0.073***	0.070***	0.061***	0.048**
			90th Percentile		
Wage Structure effect	0.076***	0.073***	0.089***	0.092***	0.076***
Reweighting Error	-0.028***	-0.022**	-0.017**	-0.024***	-0.012
Pure Wage effect	0.104***	0.095***	0.106***	0.116***	0.088***
Observations	**9,642**	**12,129**	**17,879**	**23,173**	**11,171**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.1.4

Regression of aggregate decompositions on percentile for Stay Mover (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.157***	0.124***	0.127***	0.126***	0.062*
Reweighting Error	0.017	0.003	0.028***	0.013**	0.017
Pure Wage effect	0.140***	0.121***	0.099***	0.113***	0.045
			25th Percentile		
Wage Structure effect	0.109***	0.126***	0.135***	0.106***	0.117***
Reweighting Error	0.033**	0.019	0.030***	0.015**	0.026*
Pure Wage effect	0.076***	0.107***	0.105***	0.091***	0.091***
			50th Percentile		
Wage Structure effect	0.070***	0.092***	0.082***	0.077***	0.055***
Reweighting Error	0.014	0.011	0.021***	0.013	0.028**
Pure Wage effect	0.056***	0.081***	0.061***	0.064***	0.027**
			75th Percentile		
Wage Structure effect	0.097***	0.117***	0.077***	0.055***	0.050***
Reweighting Error	0.015	0.018	0.018*	0.019	0.026*
Pure Wage effect	0.082***	0.099***	0.059***	0.036***	0.024*
			90th Percentile		
Wage Structure effect	0.143***	0.140***	0.103***	0.139***	0.087***
Reweighting Error	0.025	0.017	0.014	0.018	0.027
Pure Wage effect	0.118***	0.123***	0.089***	0.121***	0.060*
Observations	**3,055**	**4,001**	**6,221**	**8,239**	**3,189**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10^{th} – 90^{th} percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.2.1

Regression of composition effect on percentile for Move Returner (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Total Comp effect	-0.012	0.004	-0.003	-0.007	0.008
Pure Comp effect	-0.013	0.004	-0.006	-0.005	0.006
Specification Error	0.001	0.000	0.003	-0.002	0.002
			25th Percentile		
Total Comp effect	-0.003	-0.004	-0.015	-0.013	0.006
Pure Comp effect	-0.005	-0.004	-0.015	-0.012	0.004
Specification Error	0.002	0.000	0.000	-0.001	0.002
			50th Percentile		
Total Comp effect	0.004	-0.007	-0.015	-0.015	-0.003
Pure Comp effect	0.000	-0.004	-0.016	-0.010	-0.004
Specification Error	0.004	-0.003	0.001	-0.005	0.001
			75th Percentile		
Total Comp effect	-0.000	-0.005	-0.010	-0.014	-0.002
Pure Comp effect	-0.003	-0.006	-0.016	-0.01*	-0.010
Specification Error	0.003	0.001	0.006	-0.004	0.008
			90th Percentile		
Total Comp effect	-0.012	-0.013	-0.020	-0.021	-0.028
Pure Comp effect	-0.010	-0.011	-0.025	-0.024*	-0.032**
Specification Error	-0.002	-0.002	0.005	0.003	0.004
Observations	**1,400**	**1,614**	**2,736**	**3,362**	**3,922**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.2.2

Regression of composition effect on percentile for Move Stayer (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Total Comp effect	0.047*	0.035	0.032*	0.038*	0.032
Pure Comp effect	0.044***	0.032***	0.031***	0.036***	0.041**
Specification Error	0.003	0.003	0.001	0.002	-0.009
			25th Percentile		
Total Comp effect	0.044**	0.044**	0.049**	0.046***	0.044*
Pure Comp effect	0.042***	0.045***	0.048***	0.047***	0.042***
Specification Error	0.002	-0.001	0.001	-0.001	0.002
			50th Percentile		
Total Comp effect	0.050***	0.034**	0.036**	0.044***	0.053***
Pure Comp effect	0.052***	0.037***	0.038***	0.042***	0.054***
Specification Error	-0.002	-0.003	-0.002	0.002	-0.001
			75th Percentile		
Total Comp effect	0.082***	0.049***	0.052***	0.054***	0.076***
Pure Comp effect	0.084***	0.051***	0.055***	0.057***	0.075***
Specification Error	-0.002	-0.002	-0.003	-0.003	0.001
			90th Percentile		
Total Comp effect	0.107***	0.069**	0.060**	0.063***	0.095***
Pure Comp effect	0.110***	0.072***	0.060***	0.064***	0.097***
Specification Error	-0.003	-0.003	-0.000	-0.001	-0.002
Observations	**3,398**	**4,286**	**6,020**	**6,824**	**2,920**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.2.3

Regression of composition effect on percentile for NRDM (Exclude London Non-Mover)

Years	**2004/05**	**2006/07**	**2008/09**	**2010/11**	**2012/13**
			10th Percentile		
Total Comp effect	0.121***	0.071***	0.073***	0.082***	0.100***
Pure Comp effect	0.119***	0.070***	0.072***	0.081***	0.098***
Specification Error	0.01	0.00	0.001	0.001	0.002
			25th Percentile		
Total Comp effect	0.100***	0.096***	0.091***	0.092***	0.111***
Pure Comp effect	0.09***	0.094***	0.090***	0.090***	0.108***
Specification Error	0.00	0.002	0.001	0.002	0.003
			50th Percentile		
Total Comp effect	0.082***	0.053***	0.050***	0.062***	0.081***
Pure Comp effect	0.083***	0.055***	0.051***	0.065***	0.086***
Specification Error	-0.001	-0.002	-0.001	-0.003	-0.005
			75th Percentile		
Total Comp effect	0.099***	0.072***	0.060***	0.078***	0.113***
Pure Comp effect	0.101***	0.073***	0.062***	0.082***	0.116***
Specification Error	-0.002	-0.001	-0.002	-0.004	-0.003
			90th Percentile		
Total Comp effect	0.098***	0.082***	0.078***	0.097***	0.113***
Pure Comp effect	0.100***	0.083***	0.079***	0.101***	0.115***
Specification Error	-0.002	-0.001	-0.001	-0.004	-0.002
Observations	**15,116**	**18,754**	**27,400**	**34,962**	**18,666**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.2.4

Regression of composition effect on percentile for Stay Mover (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Total Comp effect	0.030	0.035	0.034	0.033	0.048
Pure Comp effect	0.030**	0.034**	0.035***	0.034***	0.049**
Specification Error	0.000	-0.001	0.001	0.001	0.001
			25th Percentile		
Total Comp effect	0.043	0.032	0.030	0.035*	0.040*
Pure Comp effect	0.043***	0.035**	0.032***	0.035***	0.041***
Specification Error	0.000	0.003	0.002	0.000	0.001
			50th Percentile		
Total Comp effect	0.040*	0.034*	0.027*	0.032**	0.043**
Pure Comp effect	0.042***	0.035***	0.029***	0.032***	0.044***
Specification Error	0.002	0.001	0.002	0.000	0.001
			75th Percentile		
Total Comp effect	0.052*	0.038	0.047*	0.050***	0.063**
Pure Comp effect	0.051***	0.035***	0.048***	0.052***	0.065***
Specification Error	-0.001	-0.003	0.001	0.002	0.002
			90th Percentile		
Total Comp effect	0.050	0.038	0.065*	0.060**	0.081**
Pure Comp effect	0.051**	0.041**	0.067***	0.059***	0.083***
Specification Error	-0.001	-0.003	0.002	-0.002	0.002
Observations	**1,942**	**2,498**	**4,084**	**5,094**	**2,702**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.3.1

Regression of wage structure effect on percentile for Move Returner (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	-0.093***	-0.124***	-0.092***	-0.088***	-0.197***
Reweighting Error	-0.012	-0.013	-0.010	-0.035***	-0.010
Pure Wage effect	-0.081**	-0.111***	-0.082***	-0.063***	-0.187***
			25th Percentile		
Wage Structure effect	-0.138***	-0.119***	-0.093***	-0.112***	-0.142***
Reweighting Error	0.010	-0.022	-0.021*	-0.030***	-0.011*
Pure Wage effect	-0.128***	-0.097***	-0.072***	-0.082***	-0.131***
			50th Percentile		
Wage Structure effect	-0.053**	-0.073***	-0.094***	-0.091***	-0.113***
Reweighting Error	-0.010	-0.022*	-0.032***	-0.030***	-0.022**
Pure Wage effect	-0.043*	-0.051**	-0.062***	-0.061***	-0.091***
			75th Percentile		
Wage Structure effect	-0.048*	-0.026	-0.054***	-0.068***	-0.087***
Reweighting Error	-0.026*	-0.013	-0.031***	-0.036***	-0.034***
Pure Wage effect	-0.022	-0.013	-0.023	-0.032*	-0.053***
			90th Percentile		
Wage Structure effect	-0.052	-0.023	-0.059**	-0.077***	-0.084***
Reweighting Error	-0.021	-0.020	-0.047***	-0.042***	-0.031**
Pure Wage effect	-0.031	-0.003	-0.012	-0.035	-0.053*
Observations	**2,784**	**3,559**	**5,547**	**7,373**	**3,799**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.3.2

Regression of wage structure effect on percentile for Move Stayer (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.076***	0.074***	0.053***	0.097***	0.054*
Reweighting Error	0.010	0.002	0.001	0.014	0.013
Pure Wage effect	0.066**	0.062***	0.052**	0.083***	0.041
			25th Percentile		
Wage Structure effect	0.044*	0.053**	0.052***	0.063***	0.056***
Reweighting Error	0.002	0.013	-0.001	-0.006	0.002
Pure Wage effect	0.032*	0.040**	0.053***	0.069***	0.054**
			50th Percentile		
Wage Structure effect	0.001	0.046**	0.028*	0.037***	0.014
Reweighting Error	-0.025*	-0.010	-0.021***	-0.022***	-0.013
Pure Wage effect	0.026	0.056***	0.049***	0.059***	0.027
			75th Percentile		
Wage Structure effect	0.006	0.027	0.026*	0.054***	-0.012
Reweighting Error	-0.031***	-0.012	-0.021***	-0.024***	-0.022*
Pure Wage effect	0.037*	0.039**	0.047***	0.078***	0.010
			90th Percentile		
Wage Structure effect	0.025	0.042*	0.064***	0.066***	0.037
Reweighting Error	-0.032**	-0.023*	-0.023**	-0.021*	-0.020
Pure Wage effect	0.057*	0.065***	0.087***	0.087***	0.057*
Observations	**3,783**	**4,895**	**7,189**	**9,104**	**3,298**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.3.3

Regression of wage structure effect on percentile for NRDM (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.012	0.050***	0.074***	0.083***	0.038
Reweighting Error	0.001	-0.006	0.001	-0.002	0.009
Pure Wage effect	0.011	0.056***	0.073***	0.085***	0.029
			25th Percentile		
Wage Structure effect	0.052***	0.053***	0.064***	0.070***	0.056***
Reweighting Error	-0.006	-0.014	0.002	-0.017**	-0.011
Pure Wage effect	0.058***	0.057***	0.066***	0.087***	0.067***
			50th Percentile		
Wage Structure effect	0.022	0.066***	0.054***	0.041***	0.029
Reweighting Error	-0.029***	-0.08**	-0.011**	-0.027***	-0.010*
Pure Wage effect	0.031***	0.074***	0.065***	0.068***	0.039**
			75th Percentile		
Wage Structure effect	0.035**	0.051***	0.053***	0.038***	0.027
Reweighting Error	-0.033***	-0.015**	-0.012***	-0.021***	-0.021**
Pure Wage effect	0.068***	0.066***	0.065***	0.059***	0.038**
			90th Percentile		
Wage Structure effect	0.077***	0.079***	0.085***	0.090***	0.074***
Reweighting Error	-0.022***	-0.013**	-0.012**	-0.026***	-0.013
Pure Wage effect	0.099***	0.092***	0.097***	0.116***	0.087***
Observations	**9,642**	**12,129**	**17,879**	**23,173**	**11,171**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

Table A.2.3.4

Regression of wage structure effect on percentile for Stay Mover (Exclude London Non-Mover)

Years	2004/05	2006/07	2008/09	2010/11	2012/13
			10th Percentile		
Wage Structure effect	0.157***	0.125***	0.124***	0.123***	0.062*
Reweighting Error	0.019	0.002	0.023***	0.012**	0.010
Pure Wage effect	0.138***	0.123***	0.101***	0.111***	0.052
			25th Percentile		
Wage Structure effect	0.102***	0.126***	0.131***	0.107***	0.115***
Reweighting Error	0.036**	0.013	0.030***	0.014**	0.024*
Pure Wage effect	0.066***	0.113***	0.101***	0.093***	0.091***
			50th Percentile		
Wage Structure effect	0.073***	0.090***	0.084***	0.077***	0.058***
Reweighting Error	0.012	0.016	0.023***	0.015	0.023**
Pure Wage effect	0.061***	0.074***	0.061***	0.062***	0.035**
			75th Percentile		
Wage Structure effect	0.093***	0.114***	0.079***	0.055***	0.052***
Reweighting Error	0.015	0.013	0.015*	0.012	0.021*
Pure Wage effect	0.078***	0.101***	0.064***	0.043***	0.041*
			90th Percentile		
Wage Structure effect	0.148***	0.147***	0.104***	0.135***	0.086***
Reweighting Error	0.024	0.012	0.013	0.012	0.024
Pure Wage effect	0.124***	0.135***	0.091***	0.123***	0.062*
Observations	**3,055**	**4,001**	**6,221**	**8,239**	**3,189**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. 10th – 90th percentile denotes the quantile of interest. All coefficient entries are rounded to three decimal places.

The table’s in the appendix illustrates the regressions that excludes London Non-Mover population when estimating.

Table’s A.2.1.1 to A.2.1.4 represents the overall effect estimations when the London Non-Mover are excluded for the four mobility groups.

Table’s A.2.2.1 to A.2.1.4 represents the composition effect estimations when the London Non-Mover are excluded for the four mobility groups.

A.2.3.1 to A.2.3.4 represents the wage structure effect estimations when the London Non-Mover are excluded for the four mobility groups.

These tables can be compared with the estimations in the main text to examine whether there is any significant difference when the population of London Non-Mover are excluded.

Chapter 3

The Impact of Horizontal and Vertical Job Mismatch on Wages: Analysis of the early career of ethnic minority Graduates in the UK

3.1. Introduction

In recent years, there has been a surge in demand for a wide array of skills, spanning both technical (STEM) and soft skills (Black et al., 2021). Policymakers worldwide have been urging individuals to enhance their STEM skills, which hold significant value in the labour market. This encouragement has prompted the proliferation of STEM programs, driven by longstanding reports highlighting the pivotal role of STEM education in fostering economic growth. Consequently, initiatives by the UK government have led to an expansion of the graduate population in the UK, steadily increasing since the 1990s (O'Leary and Sloane, 2005; Green and Zhu, 2010; Savic et al., 2019).

However, the labour market is undergoing rapid transformations in employment opportunities, necessitating graduates to possess a diverse range of 'soft' skills, including social and communication abilities, alongside expertise in their respective fields. This aspect is particularly significant as Mason et al. (2018) argue that the growing number of graduates could potentially alleviate skill shortages. These shortages, as outlined in the Industrial Strategy, have been identified as a contributing factor to the observed productivity slowdown in the UK.

This chapter delves into the issue of job mismatches in the UK, with a specific focus on recent graduates entering the labour market. Job mismatches among university graduates are categorised into two main types: vertical and horizontal. Understanding these concepts and their impact is crucial for policymakers aiming to maximise labour market efficiency and optimise the utilisation of human capital.

Vertical mismatch is also known as educational mismatch. It refers to a scenario where an individual's educational attainment does not match the educational requirements of their job. This disparity arises when a person is either overqualified or underqualified for their current position (Sloane, 2003). This mismatch can introduce inefficiencies into the job market, as individuals may find themselves either excessively qualified or inadequately prepared for their roles, potentially affecting their job satisfaction, productivity, and earnings (Mavromaras, 2012). Whereas horizontal mismatch occurs when an employee's job does not align with their field of study or expertise. Therefore, the individual may be underutilising their skill potential which leads to suboptimal allocation of human capital.

Furthermore, ethnic minority graduates in the UK face longstanding challenges in the labour market stemming from various socio-economic factors and systemic inequalities. Therefore, it is imperative to conduct further investigation into job mismatches among ethnic minorities. Currently, there is limited research examining both vertical and horizontal mismatches among ethnic minorities in the UK. To address this, an estimation will be performed to capture the impact to the earnings for different ethnic minority groups subject to being vertically and horizontally mismatched/matched.

In some cases, horizontal mismatches can be seen as an opportunity for individuals to adjust the imbalances between demand and supply when there is a lack of vacancies for specific occupations (Rudakov et al., 2019). There are two outcomes for this concept. Firstly, in the case of STEM subjects, many skills acquired are not easily transferable to other industry sectors (Boudarbat and Chernoff, 2012). Consequently, graduates in these fields are likely to experience a significant reduction in earnings if they work in occupations that result in horizontal mismatch, as they may not fully utilise their human capital potential. However, there are degree subjects where graduates are able to develop general skills such as Arts and Humanities. These skills are relatively more likely to be mismatched and the wage penalty for the mismatch is expected to be low, insignificant or even achieve a positive wage premium (Boudarbat and Chernoff, 2012).

As a result, the horizontal mismatch is complicated as the impact on the wage premium can be highly influenced differently depending on the subject choice chosen for degree subject and occupation. To understand this in depth, a multinominal logit model for subject will be estimated. This model will analyse how ethnicity influences the propensity of individuals to choose certain degree subjects compared to the White British group. The result can help researchers understand how ethnicity influences educational decision-making processes. It can identify significant predictors of subject selection across different ethnic groups and inform policies aimed at promoting diversity and equity in education.

Overall, as a baseline for vertical and horizontal mismatch, Vecchi et al. (2023) found that the overall wage penalty for purely horizontal mismatch in their dataset ranges from 1.5% to 6.4%, whereas vertical mismatch incurs a considerably greater penalty, ranging from 19.6% to 39%. However, there is a notable gap in empirical studies that simultaneously explore vertical and horizontal mismatches among ethnic minorities in the UK. Addressing this gap in the literature will provide valuable insights into understanding job mismatches and the unique challenges faced by ethnic minority graduates in the labour market.

3.2. Literature review

3.2.1. Theoretical background

Discrimination in labour markets has been extensively studied through the lens of Gary Becker's (1957) 'taste for discrimination' model. Becker's model proposes that employers, employees, or consumers may have a 'taste' or preference for not working with or hiring individuals from certain groups (such as racial or ethnic minorities). In the case of employer discrimination, firms may incur additional costs (in the form of reduced profits) to avoid hiring minority workers, even if these workers are equally or more productive than others. This behaviour stems from the employer's subjective preference, rather than any economic rationale. As a result, minority workers are either excluded from employment or relegated to lower-wage positions, despite their qualifications, leading to persistent wage gaps. Becker's model underscores the economic inefficiency of discrimination, as it distorts the allocation of human capital and reduces overall productivity in the labour market (Becker, 1957).

While Becker's model highlights the individual preferences that lead to discrimination, the theory of occupational crowding, introduced by Barbara Bergmann (1974), explores how discriminatory practices can affect labour market segmentation and the concentration of certain demographic groups in specific occupations. Occupational crowding occurs when minority groups, such as women or ethnic minorities, are systematically restricted to a limited number of job categories, leading to an oversupply of workers in those roles. This oversupply depresses wages in the crowded occupations, while underrepresentation in other fields results in wage inflation for those jobs. In this scenario, even though workers in the crowded occupations may have the skills or qualifications to perform in higher-paying roles, discriminatory barriers prevent their mobility, exacerbating wage inequalities across the labour market (Bergmann, 1974).

Both Becker's and Bergmann's models are relevant for understanding wage disparities, particularly for ethnic minorities, as discussed in this chapter. While Becker's framework focuses on individual or firm-level discriminatory preferences, occupational crowding emphasizes the structural barriers that limit occupational mobility for minority workers. These models provide critical insights into the wage effects of vertical and horizontal job mismatches for ethnic minority graduates in the UK, as explored in the empirical analysis of this chapter. By addressing both personal prejudice and structural discrimination, these theories highlight the multifaceted nature of labour market discrimination and its long-term impact on wage outcomes.

3.2.2. Background of the ethnic pay gap

The labour market position of ethnic minorities has long been a topic of debate, reflecting significant economic, political, and equality issues (EHRC, 2010; Heath et al., 2008; McGovern, 2007). There is extensive research conducted on the labour market performance of ethnic in both Europe and North America. For example, in the UK it is widely recognised that ethnic minority groups experience higher unemployment, lower earnings and reduced occupational attainment compared to Caucasians (Blackaby et al, 1998). Similarly, in the US, African Americans are the primary group experiencing labour market disadvantages, attributed in part to the spatial distribution of employment opportunities (Kain, 1968; Arnott, 1998). This divergence between residential location and job prospects in the US poses challenges for black individuals in commuting to work or accessing information about job opportunities.

Blackaby, Leslie and O'Leary (1999) found there are three focal factors that determines the unemployment for individuals with higher levels of education. Firstly, it is perceived that ethnic minorities may have a characteristic with a lower degree to its counterparts, for instance the amount of schooling acquired which directly effects the probability and position of employment. Secondly, it is deduced that for any characteristics ethnic minorities encompasses, they are penalised by a lower probability of employment as a consequence of the discrimination effect. Lastly, with the existence of discrimination for ethnic minorities, there are discrepancies to the extent and approach of the bias as well as the response of each ethnic minority.

However, many economists such as Becker (1975) argue that the primary cause of higher unemployment rates among ethnic minorities is explained by differences in individual characteristics rather than racial discrimination. Therefore, the Department of Employment Gazette in June 1995 stated that "factors which may explain the persistently higher rates of unemployment among ethnic minorities, apart from the younger age profile…include ethnic minorities generally lower level of qualifications and their industrial and regional distribution." This view of a generalised lower employment and earnings rate for ethnic minorities is subsequently recognised as a vastly partial and misleading explanation.

Leslie et al (1998) explores two other major explanations for higher unemployment and earning rates for ethnic minorities. The first alternative proposition is the effect of discrimination towards ethnic minority which is widely supported in a numerous of studies. Explicit evidence of this can be supported by the study of Brown and Gay (1985). They defined black immigrants as the treatment group and white individuals as the control group. Both groups possessed identical qualifications whereby they applied to a variety of jobs with standardised written applications. The only method that employers could distinguish the individual's ethnic background was via the applicant's name. The results are as predicted since the controlled (white) group applicants received two-thirds of positive response, whereas the treatment (black) group was subjected with considerable discrimination to the extent that employers would often falsify that the position had already been recruited. These findings are consistent with other literatures such as Smith and McIntosh (1974).

It is well known that ethnic groups experiences discrimination, although the impacts of each individual or group is often very diverse. As a result, Blackaby, Leslie and O'Leary (1999) express that due to discrimination, ethnic groups may have developed isolation attributes, which in turn adversely affects their employment and earning rates. Therefore, provided homogenous characteristics, discrimination is a key factor to understand the reason of dissimilar unemployment and earning rates across diverse ethnic groups.

Previous literatures have characterised ethnic minority workers are often poorly paid, particularly in accordance with their qualification attainment. For instance, Brynin and Longhi (2015), addresses that on average, ethnic minorities comprise higher educational qualifications in comparison to its White counterparts. Despite this fact, White majority are more likely to work in graduate occupations and subsequently achieve higher average earnings. The issue with regards to educational terms is that ethnic minorities encompass a higher probability of not being able to take full advantage of their qualification attainments. Consequently, ethnic minorities are more likely to be overqualified for their professions in comparison to their White counterparts. This factor will be extensively investigated in this research to determine whether there is a work penalty incurred by ethnic minorities in the UK in comparison to the White majority and if so, to what extent. More specifically, the ideology of work penalty will be subjugated as a wage discrepancy between the ethnic groups.

On the other hand, some academic argue it is not ethnicity itself that determines the below labour market earnings among ethnic minority groups. Conversely it is because they have different characteristics from White employees e.g., age, education, and gender. Although statistically most ethnic minority groups are less likely to be graduates compared to White employees since many of these groups are subject to poor education. Due to the various qualifications ethnic minority employees hold, this does not guarantee them a high pay employment position. Therefore, Brynin and Longhi (2015) suggests that discrimination based on ethnicity is not the fundamental reason of wage inequality.

Focus must be put upon how each ethnic minority will be affected as there will be discrepancy between the groups. There are noticeable differences with regards to the employment and inequality amongst various ethnic minority groups with the UKHLS dataset.

Brynin and Longhi (2015) found that the Bangladeshi's and Pakistani's are particularly the most vulnerable to poverty. Firstly, across all the measurements, Bangladeshi's and Pakistani's are substantially worse off than the other ethnic minority groups. From the UKHLS dataset, it is found that they have the greatest wage gap in comparison to white employees and significantly have the highest possibility to earn below the living wage. For example, there is only 16% of white male individuals that are employed below the living wage, whereas it is a staggering 57%, 39% and 28% for Bangladeshi male, Pakistani male, and Chinese male respectively. The trend is the consistent with females, although the magnitude of the discrepancy is less severe.

Moreover, the Black Caribbean group achieves a slight wage advantage over the White employees. The salary of the Black Caribbean men is around the living wage level which implies they perform the best compared to the other ethnic minorities. With regards to Black Caribbean women, they are paid significantly more than all the other ethnic minority groups and even more than White female employees. The Black African group are penalised greatly as they experience the fourth highest wage gap with regards to the ethnic minority groups. In general, the Black African have a greater likelihood to be paid less than the living wage, although the Black African women are less likely than their male counterpart. Finally, there is essentially no wage gap between the Indian and White British employees. An Indian woman is similarly paid to a White woman living wage; however, Indian men have a higher probability that they are paid less than the living wage for White men.

3.2.3. Causes of ethnic pay gap

3.2.3.1. Immigration

In this research, the individuals examined are between the working age of 16 to 64. This includes individuals that have migrated to the UK from a young age and have received education in the UK. Many ethnic minorities and especially immigrants in the UK on average receive lower wages than the average White British population. Blackaby et al (2002) and Lindley (2002) both agree that immigrants may have difficulty with the language and customs which hinders their ability within the labour market of the host country. This would reduce and impede UK immigrants to search and secure suitable jobs due to a lack of connections as well as communication skills. Often, immigrant's qualifications are not recognised in the labour market of the host country (labelled as "Other Qualification" in the research). This often leads to "occupational downgrading," where immigrants end up overqualified for their jobs, resulting in a negative wage gap (Lindley, 2009).

According to Longhi and Brynin (2017), analysing the ethnic pay gap as opposed to the immigrant pay gap becomes more complex since there are many background factors involved. As mentioned, there are two main determinants of ethnic minorities; firstly, individuals with an ethnic minority background that are born in the UK and secondly, the immigrants to the UK which also possess an ethnic minority background. In fact, in Longhi and Brynin (2017) research, they found that about half of the individuals which identify as ethnic minority are born in the UK. There are major differences between these two ethnic minority groups, the most prominent being the discrepancy of the language, standard customs and knowledge of the institutions or recognition of qualifications. British-born ethnic minorities generally possess better skills compared to their immigrant counterparts to confront these challenges, suggesting that these factors should not significantly impact their wages negatively. As hypothesised, most literatures such as Algan et al (2009) and Longhi et al (2013) conclude that the ethnic minorities born in the UK experiences a much smaller employment and negative pay gaps compared with immigrated ethnic minorities.

Despite immigrants facing numerous disadvantages in the labour market, they may possess distinct advantages compared to the majority of individuals in the UK. For instance, previous studies found that immigrants are more highly motivated as well as comprise a higher relative ability than British born individuals (Chiswick, 1980; Carliner, 1980 and Zutshi, 2008). In particular, immigrants may possess greater benefits in some circumstances than the second-generation ethnic minorities. The second-generation ethnic minorities are usually more culturally integrated within the UK society and are likely to be attributed from domestic education and from antiracial discrimination provided by the government. As a consequence, the second-generation ethnic minorities abilities and motivation is likely to be relatively reduced in comparison to their parents and ethnic minority immigrants. Further exacerbation could arise if these individuals do not receive the full benefits of domestic schooling since they are more probable to reside in areas which are less privileged than their White British counterparts (Longhi and Brynin, 2017).

Additionally, there is a positive correlation between the length of period immigrants reside in the host country and the essential knowledge, skills and connections one accumulates to match with their preferred occupation and wage according to their education level. This suggests that ethnic minorities that include inter-generation and immigrants can be complicated due to the extensive background variations. As a whole, the disadvantages that immigrants encounter is likely to be greater than the advantages they possess in the labour market as a result of discrimination, language barriers, cultural isolation, and the unrecognised overseas education.

3.2.3.2. Discrimination

Discrimination against ethnic minorities has been well identified in many previous studies such as by Lang and Lehmann (2012) as well as Guryan and Charles (2013). Discrimination of ethnic minority is usually driven by negative attitudes through social measures which can be reflected by the wage gap, particularly compared to the White British group. The wage gap is partially induced by employers assessing job applicants. For example, an employer may deduce the average attribute of an individual based on their ethnic background, this is known as the 'statistical discrimination' (Aigner and Cain, 1977). This type of discrimination is stereotyping since the employer may consider individuals from various backgrounds has different sets of work ethnics. Therefore, if the employers adhere to this approach, certain ethnic minorities are less likely to be employed or are subjected to lower wage in comparison to individuals with ethnic backgrounds that are perceived to be more productive (Longhi and Brynin, 2017). In the circumstance that all ethnic groups receive the same wage for an occupation, the ethnic minority negative wage gap is expected to exist since they would be less likely to be employed in jobs with higher salary compared to the White British group.

The existence of a wage gap between the ethnic White majority and ethnic minorities is widely acknowledged. However, since ethnic minority groups are not homogenous, the wage gap takes various forms. For example, Brynin and Güveli (2012) identified significant differences between the overall wage gap and the occupational wage gap for different ethnic groups. The former represents the average wage gap across the entire economy compared to the White British group, while the latter focuses on the average wage gap within specific occupations. The overall wage gap shows a substantial discrepancy compared to the White British group, while the occupational wage gap is considerably lower. This indicates that when ethnic minorities work in similar occupations as the White British group, the wage disparity is reduced. Consequently, the wage penalty is primarily driven by the overrepresentation of ethnic minorities in lower-paid occupations.

Previous studies have attempted to measure the extent of discrimination towards ethnic minorities within the labour market. For instance, researcher such as Wood et al (2009) utilise a common method of sending fictitious CVs to job vacancies in the labour market. The qualifications and experience listed on the CVs are identical, so the supposed productivity indicated from the applicant's form are relatively similar. The only notable difference between the applications is the candidate's name, which suggests their race or ethnicity. The underlying idea is to assess whether employers discriminate based on perceived ethnicity when considering interview opportunities.

Research consistently concludes that job applicants with foreign-sounding names statistically receive significantly fewer interview offers compared to those with native-sounding names. For example, Wood et al. (2009) conducted a study where they sent standardised applications to various job vacancies across nine occupational sectors and seven cities in the UK. As anticipated, they found that applicants with foreign names experienced a substantially lower interview acceptance rate than those with native names, despite the similar qualifications. This suggests a lower likelihood of ethnic minorities being hired compared to White British candidates. However, it remains uncertain whether this disparity directly influences the wage penalty for ethnic minorities, as indicated by Wood et al. (2009) research.

3.2.3.3. Characteristics of ethnic minority

Empirical studies utilise regression and decomposition methods to analyse the attributes of the White British group and ethnic minorities to understand the existing wage gap. Many ethnic minorities and immigrants in the UK often face challenges as their qualifications and work experience from abroad may not be recognised in the UK labour market. Consequently, these workers frequently receive wages below their productivity level due to the unfamiliar environment.

Metcalf (2009) noted the insufficient focus on worker characteristics and utilised a decomposition method to dissect the wage gap into two components. The first component, the explained part, reveals the composition effect, highlighting differences in covariates between the treatment and control groups. The second component, the unexplained part, represents the wage structure effect, indicating differences in outcomes based on these characteristics. This approach allows studies to identify the reduction of the wage gap when accounting for worker characteristics, enabling measurement of the residual wage gap as the unexplained component.

The wage structure effect, as described by Fortin, Lemieux, and Firpo (2011), is often regarded as the "sum effect of discrimination." However, the unexplained part of the wage gap should not necessarily be interpreted as the only effect of discrimination. According to Longhi and Brynin (2017), this disparity could also encompass factors such as ethnic minorities being self-employed or unemployed, contributing to the wage gap. Further explanation of the decomposition methods can be found in chapter 2 of this thesis. The estimations of the composition effect and the wage structure effect is performed and analysed.

In terms of the negative wage gap, the research by Longhi and Brynin (2017) suggest that it could stem from the ethnic minorities experiencing periods of poor economic performance while self-employed, a factor unrelated to discrimination. However, it's worth mentioning that such outcomes would not be applicable in the dataset, as it only includes full-time workers aged between 16 and 64.

To address the lack of precision, Elliot and Lindley (2008) employed decomposition techniques to analyse wage gaps separately for immigrants and various ethnic minorities. They discovered that immigrants are disproportionately represented in both below-average and above-average wage occupations, highlighting a persistent wage gap between UK-born ethnic minorities and immigrant counterparts. This suggests that personal characteristics, such as education and upbringing environment, may be primary drivers of the wage gap.

Additionally, Blackaby et al. (2002) focused on UK-born ethnic minorities to mitigate the complexities associated with immigration. Their simplified regression analysis revealed that individual characteristics could statistically explain about half of the wage gap between Black and White ethnic groups in the UK, primarily driven by differences in educational qualifications. Interestingly, they did not find significant explanatory power for the wage gap among Pakistani or Indian ethnic groups. Thus, due to the heterogeneity among ethnic groups, further decomposition is necessary, as increased aggregation levels make it challenging to use characteristics to explain the wage gap (Longhi et al., 2012).

3.2.4. Job mismatch – Ethnic minority

In the labour market, there are two types of job mismatches: vertical and horizontal. Vertical mismatch arises from misalignment between an individual's qualifications and the demands of their job, leading to overeducation or undereducation. Horizontal mismatch occurs when skills or qualifications do not match job requirements.

3.2.4.1. Vertical mismatch

Educational vertical mismatch occurs when an individual's level of education does not match the educational attainment required for the specific occupation. This discrepancy occurs when an individual is either overeducated or undereducated with respect to their occupation. Overeducation happens when an individual’s level of education is greater than what is required for their job and vice versa for undereducation (Sloane et al, 1999). In circumstances of vertical mismatch, individuals typically receive a lower (higher) wage relative to their level of educational attainment if they are overeducated (undereducated) for their position.

Several studies have examined the wage gap resulting from vertical mismatch and its impact on ethnic minorities. However, there is a scarcity of research exploring the wage gap for ethnic minorities in conjunction with vertical mismatch. It is widely acknowledged that a significant portion of the UK labour force experiences overeducation or undereducation. Sloane et al. (1999) found that approximately 31% of British workers were overeducated and 17% were undereducated, leaving the remaining workers educationally well matched for their occupations.

Various methods exist to assess whether individuals are overeducated, undereducated, or well-educated. McGuinness (2006) employed a subjective approach, asking respondents to compare the minimum educational requirement for their job with their actual education level. Alternatively, Dolton and Vignoles (2000) analysed data from the 1980 National Survey of Graduates and Diplomates, defining overeducation as graduates working in jobs requiring sub-degree qualifications. They found that 38% of graduates were overeducated initially, declining to 30% after six years. Dolton and Vignoles (2000) noted that overeducated graduates earned substantially less than those in graduate-level roles.

Objective measures, as suggested by Kupets (2016), include identifying the average or modal education level for an occupation or comparing educational requirements using classifications such as ISCO or DOT. In the study, the modal method was employed, defining individuals with higher (lower) education than the modal level for their occupation as overeducated (undereducated) (Kiker et al., 1997). Rumberger (1987) used the DOT to convert occupational education requirements into equivalent years of education, comparing them with employees' actual education. Additionally, the mean-based statistical approach utilizes dataset variability to determine the required education level for an occupation. An overeducated (undereducated) individual has education levels one standard deviation above (below) the mean for their occupation (Velasco, 2021).

Overall, existing empirical findings consistently demonstrate that individuals with higher educational attainment tend to receive higher wages. However, issues arise when there is an excess of education relative to job requirements. While most overeducated workers still earn a positive return on their education, a significant proportion experience a wage penalty (Alba Ramirez, 1993; Hartog, 2000; Sloane et al., 1999). This penalty reflects the wage difference between a perfectly matched worker and one who is overeducated for their job (Tsang et al., 1991). Overeducated workers may feel dissatisfied with the disparity between their actual and required education, leading to lower job satisfaction and reduced productivity, thus decreasing their returns on education (Battu et al., 2000).

Battu and Sloane (2002) provide insights into vertical mismatch for non-white individuals in the UK. Their study, based on the Fourth National Survey of Ethnic Minorities (FNSEM), reveals a notable bias toward overeducation among ethnic minorities compared to the White British group. When considering foreign qualifications, all ethnic minority groups exhibit increased levels of overeducation, exceeding 30%. Conversely, undereducation levels are lower for ethnic minority groups relative to the White British group. Additionally, Battu and Sloane find that the African-Asian group is significantly more prone to overeducation compared to Indians. Factors such as possessing foreign qualifications, being born in the UK, and fluency in the local language increase the likelihood of both overeducation and undereducation.

Battu and Sloane (2002) confirm the general hypothesis of a positive correlation between surplus education and earnings. Their estimation reveals increase in earnings for individuals with education well-matched to their occupation, and a negative return for those who are undereducated. These findings are particularly pronounced for ethnic minorities. Interestingly, the longer ethnic minorities reside in the UK, the greater their access to private transport, leading to increased earnings. Additionally, there is a positive correlation between the concentration of a specific ethnic minority group and their earnings. This phenomenon may be attributed to employer’s recognition and appreciation of foreign qualifications which results in an increase in heterogeneity and productivity levels within the workforce.

3.2.4.2. Horizontal mismatch

The investigation of horizontal mismatches is scarce compared to vertical mismatches, primarily due to an insufficient reliable data on university graduates. Additionally, the absence of self-evaluation data on job-education alignment contributes to this gap in research (Gimpelson et al., 2010). Nevertheless, horizontal mismatch is able to instigate a greater impact on the wage premium than vertical mismatch of university graduates. The causes and consequences of job-education mismatches concerning graduate salaries can be examined through various labour market theories such as the human capital (Mincer, 1974), job matching (Jovanovich, 1979), and job assignment (Sattinger, 1993). However, there are no previous studies found which simultaneously research horizontal mismatch and ethnic minority. Therefore, previous literature can only be assessed solely on horizontal mismatch.

Robst (2006) hypothesised educational mismatch is more prevalent among graduates who have pursued degree fields that offer general skills, such as Arts and Humanities subjects, as opposed to graduates with degrees that provide specific human capital, such as STEM subjects. The issue of job-education mismatches is intricately linked to degree subject choices. Individuals invest resources and time into acquiring skills in a specific field of study, expecting employment opportunities aligned with their education. When job-education mismatch occurs, inefficiencies arises and wage penalties for workers. The direction of this impact hinges on whether the mismatch is supply-related, driven by individual choices, or demand-related, influenced by factors like job availability and market competition (Nordin et al., 2010).

A mismatch driven by demand factors is expected to have an adverse effect on earnings. This can occur due to a scarcity of demand for specific fields in the economy or adverse selection among mismatched individuals, where those involuntarily mismatched may have lower abilities compared to their counterparts. This instigates potential endogeneity issues in empirical analyses of the impact of job-education mismatch on earnings, as graduates with lower abilities are more likely to experience mismatch and consequently earn less. Conversely, if the negative impact of job-education mismatch solely reflects lower abilities, its effect should be consistent across different fields. Moreover, if the mismatch is supply-related, it is less likely to be influenced by sorting based on abilities (Robst, 2007).

The influence of supply-related mismatch is uncertain or even positive since it stems from the graduate's own decision and is influenced by changes in their preferences after attending university. An explanation could be their improved understanding of job characteristics within their chosen field (Nordin et al., 2010). According to job matching theory, recent graduates often lack understanding of how their skills and abilities align with job requirements (Jovanovich, 1979). Consequently, they attempt to search the optimal fit relative to their skills and occupation. This increases job mobility and voluntary mismatches between education and employment. This is supported by the job assignment theory, which suggests that the return on education is contingent on the effectiveness of matching heterogenous graduates with diverse occupations (Sattinger, 1993).

3.3. Data

3.3.1. Office for National Statistics (ONS) – Labour Force Survey (LFS)

(ONS) Quarterly Labour Force Survey (QLFS), which serves as a unique source of information on economically active individuals within the UK. Economically active individuals include both employed and unemployed people who are actively participating in the labour force.

This dataset covers a broad array of topics including occupation, working hours, level of education, and other personal characteristics of household members aged 16 and above. This research examines the dataset by conducting a pooled cross-section between the years 2012 and 2020. Conducted quarterly since 1992, the survey provides a comprehensive overview of ethnic minority groups, thereby enhancing the dataset's utility. With this data, the UK labour market performance of ethnic minorities can be compared to the White British group.

Lindley (2007) highlights the challenge of analysing ethnic minority characteristics due to limited observations. However, the sample size of the QLFS enables a detailed examination of ethnic minority groups. This study will focus on key information such as earnings, employment status, tenure, and various socio-economic factors including age, gender, marital status, and human capital indicators such as years of education and highest qualification attained. This sampling approach offers an advantage by oversampling ethnic minority groups and acknowledging the diversity within the non-white population.

There are a total of nine categories of ethnic groups: Black Caribbean, Black African, Bangladeshi/Pakistani, Chinese, Indian, Other Asian background, White British, Other White, and Other Ethnic groups. Bangladeshis and Pakistanis are combined due to similar characteristics and relatively lower respondent numbers compared to other ethnic groups (Lindley, 2007). Additionally, Blacks are subdivided into three groups (Jackson et al., 2011) to account for significant discrepancies in characteristics such as highest qualifications held and earnings outcomes.

The period between 2012 and 2020 is specifically selected for the cross-section analysis due to the sufficient sample size of minority ethnic groups. Including the White ethnic group, there is a total of 261,814 observations. Additionally, the period between 2007 and 2008 follows the aftermath of the global financial crisis and with the aftereffects still remaining in 2011 (Gopinath, 2020). To mitigate the potential bias introduced by this unprecedented event, the analysis concentrates on data from 2012 onwards.

3.3.2. Descriptive statistics

The descriptive statistics exclude respondents who are unemployed, not within the working age range of 16 to 64, and individuals who are self-employed, as earnings information is not available for this group. Additionally, respondents with “Missing Data” or “Does not Apply” for the following variables are discarded: Full or Part-Time status, Ethnic Group, Hours Worked, Gross Hourly Pay, Age upon Completion of Education, and Highest Qualification Held, as these are essential pieces of information required for analysis.

Table 3.1 presents the return with respect to the duration of education, proportion in employment, real hourly earnings and the length of tenure. In column 1, the White British ethnic group exhibits the lowest average years of education compared to its counterparts, with only an average of 18.21 years. Following closely, the Black Caribbean group experiences an average of 18.52 years of education. The Chinese ethnic group records the highest years of education at 21.45 years, which aligns with the common perception of Chinese individuals being high achievers. The remaining ethnic minority groups demonstrate figures similar to each other with a range between 19.85 to 20.64.

It is interesting that despite both the Black African and Black Caribbean ethnicities having ancestral ties to Africa, there exists a notable education gap of 2.12 years. These findings suggest that all ethnic minority groups possess a higher number of years of education compared to the White British ethnic group.

Column 2 of Table 3.1 displays employment rates across ethnic groups. Despite the White British group having the lowest education levels, they surprisingly have the highest employment rate at 87.30%. This coincides with Zwysen et al. (2021) findings in LSE British Politics and Policy, indicating widespread discrimination against ethnic minorities in hiring. They highlight that ethnic minorities experience discrimination which lead to penalties in employment outcomes. Discriminated groups experience unequal treatment in the labour market, emphasising the need for systemic changes to promote fair employment opportunities.

The Black Caribbean and Indian group also exhibit high employment rates of 82.77% and 82.07%, respectively. Notably, despite these groups achieving relatively high employment rates, the Black Caribbean group has the second-lowest years of education, while the Indian group has the third highest years of education. These findings are consistent with the Annual Population Survey (APS) conducted by the ONS, where the White British group demonstrates the highest employment rate, closely followed by the Indian group.

Furthermore, research such as the Catney et al. (2016) draws on evidence from the 1991 to 2011 UK census that among economically active individuals, only Indian men and Black Caribbean women achieve similar employment rates to the White ethnic group.

Table 3.1

Means of selected variables

Variables	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Pakistani and Bangladeshi	White British
Years of Education (Years)	20.64	18.52	21.45	20.61	20.14	20.47	19.85	18.21
Employed (%)	74.66	82.77	71.92	82.07	76.52	71.45	62.19	87.30
Real hourly earnings (£/hour)	12.44	13.26	16.23	15.81	12.78	13.67	12.00	14.18
Tenure (Years)	4.86	5.40	5.05	5.31	5.16	4.92	4.96	5.58

There is an employment discrepancy between the Black Caribbean and Black African ethnic groups, despite both categorised as Black ethnicity. For instance, the Black African group demonstrates an employment rate of 74.66%, which is considerably lower than the Black Caribbean group by 8.11%. Interestingly, the employment rate of the Black Caribbean group surpasses the Black African group, even though the latter has a higher average level of education. This justifies the importance of not grouping ethnic minorities together solely based on perceived similarities in background.

Evidence from the Noronha (2022) corroborates these findings, revealing that African, Other Black, and Black African groups experience high levels of unemployment, ranging from 13% to 17%, although rates were comparatively lower for Caribbean women. This highlights the importance of recognising the diverse experiences and challenges faced by different ethnic subgroups within broader classifications.

Table 3.1 indicates that, apart from the Pakistani and Bangladeshi groups, other ethnic groups such as the Chinese, Other Asian Background, and Other Ethnic Background exhibit similar employment rates. Specifically, the Chinese group has an employment rate of 71.92%, the Other Asian Background group has a rate of 76.52%, and the Other Ethnic Background group has a rate of 71.45%.

Interestingly, despite the Chinese ethnic group participating in education for the highest number of years on average, their employment rate is relatively low compared to other groups. On the other hand, the White British group, which participates in the lowest number of years of education, has the highest employment rate among all the groups analysed. This observation suggests that there is not a direct correlation between the number of years of education and the employment rate. Other factors, such as discrimination, skill match, economic conditions, and cultural factors, might play significant roles in determining employment outcomes. Therefore, simply increasing education levels may not necessarily lead to higher employment rates, and a more understanding of the factor's influencing employment is required.

In line with previous studies, the findings indicate that the combined Bangladeshi and Pakistani ethnic group has the lowest employment rate, standing at 62.19%. The trend of ethnic minority groups generally exhibits a lower employment rate compared to the White British majority. Li (2020) illustrates that this discrepancy is largely attributed to the significantly lower employment rates experienced by these ethnic minority groups in 2001. For instance, data from the General Household Survey and the Labour Force Survey reveal a substantial increase of 20.6 percentage points in the employment rate of the Bangladeshi ethnic group between 2001 and 2019, while the White ethnic group saw a mere 4.0 percentage point increase over the same period. Hence, while the employment rates of ethnic minority groups compared to the White British group are reducing, there persists a substantial employment gap.

The analysis of hourly earnings reveals that the Chinese and Indian ethnic groups significantly outearn other ethnic minority groups, with real hourly earnings of £16.23 and £15.81 respectively. Remarkably, both the Chinese and Indian ethnic groups earn more than the White British group, with real hourly earnings of £14.18. This aligns with findings from the Office for National Statistics (ONS, 2020), which focused on 2019 data and indicated that the Chinese and Indian ethnic groups earned more than the White British by 23.1% and 15.5% respectively.

In general, there appears to be a positive correlation between years of education and hourly earnings. This is evident in the case of the Chinese and Indian ethnic groups, which boast the highest average real hourly earnings, largely attributed to their relatively higher years of education. However, there are exceptions to this trend. For instance, despite having a higher average number of years of education (20.64 years compared to 20.61 years for the Indian ethnic group), the Black African ethnic group has lower real hourly earnings (£12.44) than the Indian ethnic group (£15.81). Similarly, the Other Asian Background group and the Other Ethnic group possess similar years of education to the Indian group, yet they earn considerably less, with real hourly earnings of £12.78 and £13.67 respectively.

This suggests that other unobserved factors are likely to significantly influence hourly earnings, such as family background and region of workplace. These factors may contribute to discrepancies in earnings despite similar levels of education, highlighting the complexity of the relationship between education and earnings.

Additionally, the Black Caribbean ethnic group's real hourly earnings are 6.59% higher than those of the Black African ethnic group, despite the Black African group having a higher level of education by 11.45%. Clerk (2022) investigated the median hourly earnings in England and Wales in 2019 and similarly found that the Black Caribbean group earned more than the Black African group, with figures of £12.03 and £11.50 respectively.

Lastly, it is evident that the Bangladeshi and Pakistani ethnic group consistently performs the poorest across various aspects in comparison to its counterparts. This is particularly evident in terms of real hourly earnings, where the Bangladeshi and Pakistani group exhibits the lowest figure at £12.00. Consequently, the Bangladeshi and Pakistani group experiences the lowest employment rate as well as the lowest real hourly earnings compared to all other ethnic groups.

Table 3.2 presents a summary of the educational attainment of various ethnic groups, recording the results of the highest qualification achieved. Six education dummy categories are generated, ranked from highest to lowest attainment as follows: degree (including both undergraduate and postgraduate levels), higher education, A-Levels, GCSE, other qualification, and no qualification. The ethnic group White British is used as the focal baseline, given its status as the ethnic majority in the UK (comprising 222,913 observations). Consequently, it will primarily serve as a reference point for comparing other ethnic minority groups (aggregating 36,369 observations).

When it comes to educational qualifications, compared to its ethnic minority counterparts, the White British group exhibits the highest percentage of degree holders at 32.32% and the lowest percentage of individuals without qualifications at 4.11%. Additionally, the White British group also achieves the highest percentage of A-levels and GCSEs, with figures of 24.27% and 23.11%, respectively. These disparities may stem from the relatively greater financial stability of White British households compared to ethnic minority households. Consequently, White British individuals may have the means to pursue further education without needing to contribute financially to their households, whereas many ethnic minority households rely on additional financial support from their teenage offspring.

Surprisingly, the overall representation of White British group possesses a relatively low population count of degree holders with respect to the other minority ethnic groups. In the case of degree holders, the Black Caribbean group is an anomaly as it is the only ethnic minority group which has a lower population percentage count (30.43%) in comparison with the White British group (32.51%). This finding is consistent with previous studies as the Institute for Fiscal Studies (IFS, 2015) has also found that on average, White British individuals are the least likely ethnic group in the UK to attend university. Crawford and Greaves also state that Black Caribbean individuals are often the most under-represented group that attend higher education which is coherent with the result.

Despite being the majority ethnic group, the White British group demonstrates a relatively lower proportion of degree holders compared to other minority ethnic groups. Notably, the Black Caribbean group stands out as an anomaly, with a lower population of degree holder than the White British group with figures of 30.43% and 32.51% respectively. This is consistent with findings from previous studies (Crawford et al., 2015), which indicate that White British individuals are generally less likely to pursue higher education in the UK.

The educational attainment of Indian and Chinese ethnic groups exceeds their White British counterparts. For instance, nearly double the proportion of Indians (57.53%) achieve degrees compared to White British (32.51%) individuals. Additionally, the Chinese group demonstrates the highest percentage of degree holders, surpassing even the Indian group. Table 3.2 shows that 70.53% of the Chinese minority population holds a degree, which is 38.02% higher than their White British counterparts. This suggests that the number of Chinese degree holders is more than double that of White British individuals. Zwysen and Longhi (2016) have observed that students from Chinese and Indian ethnic backgrounds are more likely to graduate from prestigious universities, resulting in better labour market outcomes for these minority groups.

In line with findings by Shiner and Modood (2002), Table 3.2 indicate Black Caribbean individuals tend to have the lowest level of degree attainment compared to other ethnic groups. However, the results diverge when it comes to the Black African group. Despite sharing similar backgrounds, the dataset reveals a significant disparity in degree attainment between Black African (47.83%) and Black Caribbean (30.43%) individuals.

Table 3.2

Highest qualifications across ethnic groups (%)

Highest Qualification	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Other White	Pakistani & Bangladeshi	White British
Degree	46.66	33.36	68.48	55.19	47.23	46.2	50.49	45.35	33.40
Higher education	13.20	15.16	5.01	9.05	8.02	7.88	8.63	6.79	10.58
A-Levels	12.44	20.84	6.92	11.41	12.63	10.15	10.78	19.05	21.53
GCSE	9.96	19.55	7.51	10.45	8.02	7.26	5.83	15.21	25.19
Other Qualification	12.79	7.04	8.84	10.07	18.44	19.41	19.50	8.94	5.26
No Qualification	4.95	4.06	3.24	3.82	5.68	6.87	4.77	4.65	4.03
Observations	**1,978**	**1,478**	**679**	**3,138**	**1,497**	**1,726**	**8,368**	**1,354**	**113,520**

The House of Commons Library (2020) analysis of the educational outcomes of Black pupils and students reveals a 28% rise in the university entry rate of Black young individuals in 2010, representing the most significant increase among ethnic groups. This trend is supported by the dataset as it indicates a substantial increase in the percentage of Black African individuals attaining degrees, surpassing those of the White British group by 15.32%.

Despite the increased enrolment of Black ethnic minorities in universities, research conducted by Roberts and Bolton (2020) suggests that they are less likely to achieve high grades, gain admission to prestigious universities, or secure high-skilled jobs. This assertion is supported by UCAS admissions data from 2019, which indicates that 40,000 Black African students were admitted for undergraduate study in England, whereas only 8,600 Black Caribbean students were admitted. These findings underscore the disparities in educational outcomes among different Black ethnic groups. The degree holders in the remaining groups, Other White at 46.76% and Other Asian Background at 45.28%, reflect a comparable proportion to that of the Pakistani, Bangladeshi, and Black African groups.

Regarding Higher Education qualifications, the dataset shows relatively low attainment percentages across ethnic groups, ranging from 4.72% for the Chinese minority to 12.96% for the Black Caribbean minority. Analysis of Table 3.2 (Column 2) suggests an inverse relationship between Higher Education attainment and Degree outcomes, except for both Black ethnic groups, which exhibit higher results compared to others. This is consistent with Roberts and Bolton's (2020) findings that young people from Black ethnic backgrounds are more likely to pursue higher education. However, despite the relatively high number of Black ethnic individuals obtaining Higher Education qualifications, retention rates tend to be lower and grade outcomes poorer than average (Roberts and Bolton, 2020).

The results in the A-Level category reveal an inverse relationship compared to the first two columns. For instance, the White British group represents the highest proportion (24.19%) of individuals attaining A-Levels as their highest qualification, followed by the Black Caribbean group at 21.97%. Conversely, the Chinese ethnic minority group has the lowest representation in the A-Level category, with only 6.29%. These findings align with expectations, as ethnic minority groups (except for the Black Caribbean group) tend to pursue further studies to attain a degree, contrasting with the choices of the White British group.

According to the Ethnicity facts and figures in 2020, White British experiences the lowest percentage of students going into sustained education after the age of 18 in comparison to all the other ethnic minority groups. Department for Education in 2020 findings record that between 2017 to 2018, the White British apprenticeship and employment rate are the highest with figures of 11% and 28% respectively (other category includes education, no sustained education/employment and unknown). This outcome instigates the low level of White British group to progress to further education after studying A-Levels which is consistent to the findings. Other ethnic group results from the Ethnicity facts and figures are also consistent to the findings. For instance, the Chinese is by far the highest group which continues in higher education studies after A-Levels with a figure of 79%, while the apprenticeship and employment resulting to be low with figures of 2% and 6% respectively. With regards to the GCSE category, it follows a similar pattern to the A-Level classification.

According to Ethnicity Facts and Figures (2020), the White British demographic exhibits the lowest percentage of students continuing into sustained education after the age of 18 compared to other ethnic minority groups. Additionally, findings from the Department for Education (2020) reveal that between 2017 and 2018, the White British group boasted the highest apprenticeship and employment rates, at 11% and 28% respectively. This suggests a lower inclination for the White British group to pursue further education after completing A-Levels, aligning to the findings.

Moreover, the Chinese ethnic group demonstrates the highest rate of progression into higher education studies after A-Levels, standing at 79%, consistent with the findings from Ethnicity Facts and Figures. However, their rates for apprenticeship and employment are comparatively lower, at 2% and 6% respectively. Similarly, the pattern observed in the GCSE category mirrors that of the A-Level classification.

Except for the Other White and Other Asian Background groups, all ethnic groups exhibit a low percentage population for the "other qualification" category. For instance, the range extends from 6.21% for White British to 13.92% for Black African, while the populations for

Other White and Other Asian Background in the "other qualification" category are notably higher at 22.05% and 19.17% respectively. Interestingly, these two groups consistently display homogeneous results across various highest qualification categories. This may suggest that the Other White and Other Asian Background ethnic groups share marginally corresponding characteristics.

Finally, across all ethnic groups, there is a relatively low level of individuals classified as "no qualification". The Pakistani and Bangladeshi group has the highest percentage, standing at 8.29%, while the Chinese group has the lowest, with only 4.06%. These results align with Battu and Sloane's research, which identified Bangladeshis and Pakistanis as having the highest incidence of no qualifications among ethnic groups. Additionally, their study noted that the Chinese group has a notably lower rate of no qualifications compared to all other ethnic groups, providing further support for the findings.

Overall, the findings provide evidence supporting the presence of White British prejudice or an ethnic minority penalty regarding employment rates and real hourly earnings. Further incidence of mismatch can be examined with other explanatory variables such as 'managerial statuses of ethnic groups' (detailed statistics in Appendix A.3.1). For example, while the White British group has the second-lowest degree attainment rate, they hold the highest rate of 'Manager' positions among all ethnic groups. With a managerial rate of 28.44%, the White British group surpasses even the highest achievers among ethnic groups, such as the Chinese (25.60%) and Indian (26.68%) ethnic groups, which have degree attainment rates of 70.65% and 57.40% respectively. Consequently, there is a degree of overeducation among ethnic minority groups, particularly when compared to the White British majority group.

Tables 3.3 and 3.4 display the highest qualifications of each ethnic group for males and females, respectively. While the first column (degree holders) generally aligns for both genders, notable discrepancies appear among several ethnic minority groups. For example, high-attaining groups like the Chinese and Indian show a higher percentage of males graduating from university compared to females. Specifically, the Chinese and Indian male groups exhibit 5.07% and 4.64% higher proportions, respectively, than their female counterparts. This phenomenon may stem from cultural perspectives passed down through generations, where there is traditionally an expectation for males to pursue higher education, while females are encouraged to focus on household responsibilities. Consequently, this cultural norm contributes to a disparity favouring male attainment, resulting in a higher percentage of degree holders among Chinese and Indian males compared to females, on average.

Interestingly, Black African and Black Caribbean individuals, both originating from black backgrounds, exhibit contrasting outcomes in terms of gender. For instance, the male population of the Black African group (49.49%) exceeds that of females (46.66%), whereas the male population of the Black Caribbean group (24.78%) is lower than their female counterparts (33.36%). According to Heath and Cheung (2006), black African men are among the most likely to possess degree-level or higher qualifications. In contrast, Healy et al. (2007) found that black Caribbean females are disproportionately represented in health and social care, potentially justifying the significantly higher prevalence of degree-level qualifications among black Caribbean females compared to males.

Overall, there is not much difference between genders for both A-Level and GCSE qualifications. However, it is notable that the Bangladeshi and Pakistani female ethnic groups significantly outnumber their male counterparts. Female Bangladeshi and Pakistani students have an A-Level and GCSE surplus of 6.84% and 3.54% respectively compared to males. Initially surprising, this finding can be attributed to the historical struggles of Bangladeshi and Pakistani females to overcome barriers to education, especially during the first half of the 1990s (Takhar, 2016). Traditionally, parents from these ethnic backgrounds-imposed restrictions on female education (Bhopal, 2010; Hussain and Bagguley, 2007), leading to a racialized gender group that faced discrimination and beliefs hindering academic achievement at A-Level and GCSE (Basit, 1997; Mirza, 2006).

Despite these constraints, recent trends show that Bangladeshi and Pakistani females have been better represented than young British Bangladeshi males, as discovered by Niven et al. (2013). This shift may be attributed to their active engagement in securing a better future. Takhar (2016) describes this as the operationalisation of agency and 'agentic autonomy,' highlighting how a generation of females is actively involved in social transformation, leading to educational improvements and reshaping future gender relations.

Table 3.3

Highest qualifications across ethnic groups (%) for male

Highest Qualification	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Other White	Pakistani & Bangladeshi	White British
Degree	49.49	24.78	73.55	59.83	43.29	49.6	42.79	45.53	31.63
Higher Education	11.26	9.15	4.69	6.38	7.77	7.62	7.72	5.68	9.22
A-Levels	10.80	23.77	5.63	9.23	12.37	9.23	12.18	12.16	26.97
GCSE	8.80	24.11	4.69	9.29	8.82	8.06	6.42	11.67	20.84
Other Qualification	15.43	12.83	7.32	10.23	19.61	17.2	24.79	14.49	7.17
No Qualification	4.22	5.36	4.12	5.04	8.14	8.29	6.10	10.47	4.17
Observations	**109,227**	**1,750**	**896**	**533**	**3,607**	**1,326**	**1,479**	**7,589**	**2,236**

Table 3.4

Highest qualifications across ethnic groups for female (%)

Educational Mismatch	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Other White	Pakistani & Bangladeshi	White British
Over-Educated	80.59	65.16	84.98	80.11	81.9	82.51	89.01	70.75	56.62
Under-Educated	7.48	9.20	5.30	5.16	7.62	7.17	4.23	6.28	11.54
Well-Matched	11.93	25.64	9.72	14.72	10.49	10.32	6.76	22.97	31.83
Other Qualification	12.79	7.04	8.84	10.07	18.44	19.41	19.5	8.94	5.26
No Qualification	4.95	4.06	3.24	3.82	5.68	6.87	4.77	4.65	4.03
Observations	**1,978**	**1,478**	**679**	**3,138**	**1,497**	**1,726**	**8,368**	**1,354**	**113,520**

It is evident that the "No qualification" category in Table 3.2, Table 3.3 (male), and Table 3.4 (female) exhibits a relatively low percentage compared to other highest qualification categories. However, Bangladeshi and Pakistani males stand out as an anomaly, with a significantly higher percentage of the population lacking qualifications compared to females. Specifically, females have a figure of 4.65%, while males have 10.47%, indicating that there are over twice as many males employed without any qualifications compared to females.

Since the compulsory implementation of GCSE in September 1987 (Brooks, 2014), respondents lacking any qualification necessarily imply they are either at least middle-aged or migrants to the UK from overseas. This finding aligns with research from the Centre on Dynamics of Ethnicity (CoDE, 2014), which found that among Bangladeshi and Pakistani migrants, 31% and 28% respectively aged 25 to 49 have no qualifications. Moreover, CoDE recorded that over half of Bangladeshi and Pakistani migrants aged 50 to 64 lack any qualifications, the highest among all ethnic groups. Similar results were found by Focus on Social Inequalities (2004) using ONS data, indicating that both Bangladeshi and Pakistani men and women are more likely to lack qualifications compared to other ethnic groups.

A contributing factor to the higher percentage of no qualifications among Bangladeshi and Pakistani males is their comparatively lower socio-economic status, as highlighted by the Trades Union Congress (TUC, 2006), which states, "The position of people of Pakistani and Bangladeshi origin is vitally important for all anti-poverty campaigners because they are far more likely to be poor than any other ethnic group." Consequently, males from these backgrounds are often expected to start full-time work at a young age to provide financial support to their families, limiting their opportunities for education and qualification attainment.

Conversely, a significant proportion of Bangladeshi and Pakistani females are employed in hospitality and care occupations, which often require some form of qualification. These factors underscore the considerable discrepancy in qualification levels between Bangladeshi and Pakistani males and females.

In this study, the concept of educational mismatch is introduced to categorise whether respondents' education aligns well with their occupations. To determine this outcome, the modal method is employed, an alternative to the mean approach adapted from Battu and Sloane (2002). The modal method identifies the most common level of education required for each specific occupation. As depicted in Table 4, three educational mismatch outcomes are considered: under-educated, adequately educated (well-matched), and over-educated.

A respondent is classified as under-educated if their educational level falls below the modal value required for their occupation. Conversely, individuals are considered well-matched if their educational level matches the modal requirement for their occupation, and over-educated if their education level exceeds what is typically required for their occupation.

The modal method is advantageous as it is less susceptible to outliers and technological or workplace changes (Kiker et al., 1997). Occupations are categorized using the 2-digit SOC level, and in cases where occupations have fewer than 10 observations, they are merged with suitable and adjoining occupations. Subsequently, for each ethnic group, the modal level of education is determined to assess whether respondents are mismatched or suitably matched to their occupations.

In Table 3.5, column 2, it is evident that both the White British and Black Caribbean groups are significantly more well-matched in terms of education compared to other ethnic counterparts, with levels of 33.86% and 29.42% respectively. This observation aligns with expectations as the White British group constitutes over 85% of the overall observations, introducing a bias in favour of this group due to the modal method adopted from Battu and Sloane (2002). While this systematic result is beneficial for using the White British group as a benchmark or treatment effect, it differs from findings by Battu and Sloane (2002), who reported that Whites exhibited the lowest percentage of well-matched individuals compared to all other ethnic minority groups, with a figure of 41.59%.

Furthermore, Battu and Sloane report a significantly higher proportion of respondents classified as adequately educated, with every ethnic group having higher figures compared to the results. This difference may be attributed to the increased number of individuals opting for higher education studies in the 21st century compared to the 20th century. This trend is supported by Coughlan (2019), who revealed figures from the Department for Education indicating that over 50% of young individuals are now attending university. In

Table 3.5

Educational mismatch (%)

Educational Mismatch	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Other White	Pakistani & Bangladeshi	White British
Over-Educated	82.35	61.28	85.02	80.59	80.86	80.02	88.08	71.84	54.15
Under-Educated	6.20	9.30	5.30	5.37	7.64	7.52	4.57	7.06	11.99
Well-Matched	11.45	29.42	9.69	14.04	11.50	12.46	7.35	21.1	33.86
Observations	**3,728**	**2,376**	**1,208**	**6,722**	**2,827**	**3,205**	**15,939**	**3,569**	**222,913**

contrast, in 1990, only 25% pursued all forms of post-18 education, and in 1980, just 15% remained in full-time education after the age of 18. Consequently, this transformation explains the substantial shift in the proportion of educational mismatches between Battu and Sloane's study in 2002 and the updated dataset.

With the exception of the White British and Black Caribbean groups, the other ethnic groups demonstrate considerable mismatches in terms of educational levels. The lowest incidence of well-matched individuals is observed among the Other White group at 7.35%, while the high-performing Asian groups also exhibit low proportions of well-matched individuals: Chinese (9.69%), Indian (14.04%), Other Asian Background (11.5%), and Other Ethnic Group (12.46%). Additionally, the Black African group (11.45%) significantly differs from the Black Caribbean group by 17.97%. Finally, the Bangladeshi and Pakistani group is positioned in the midpoint with 21.1%.

As previously stated, the findings show an increase in the over-educated category compared to Battu and Sloane (2002). This trend is likely due to the exponential increase in the number of young individuals pursuing higher education studies.

Tables 3.6 and 3.7 illustrate the educational mismatch among males and females, respectively. Overall, there is not a significant difference between males and females within each ethnic group across the three educational mismatch categories. However, Black African males exhibit a lower percentage of under-educated individuals compared to females, with a difference of 2.62%. Also, in this group females comprise a lower percentage of over-educated individuals compared to males, with a difference of 3.75%. This suggests that, on average, Black African males are less likely to be under-educated but more likely to be over-educated relative to their occupations compared to Black African females.

Despite both having African ancestral origins, there are extensive differences in the educational mismatch configuration between the Black African and Black Caribbean groups. Overall, the Black African group tends to fare better than their Black Caribbean counterparts in terms of a substantially higher percentage of over-educated individuals. However, the Black Caribbean group consists of significantly more well-matched individuals than the Black African group. Interestingly, within the Black Caribbean group, there is a notable disparity between males and females in both the well-matched and over-educated categories.

Table 3.6

Educational mismatch for male (%)

Educational Mismatch	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Other White	Pakistani & Bangladeshi	White British
Over-Educated	84.34	54.91	84.99	81.04	79.79	78.26	86.80	72.63	51.56
Under-Educated	4.86	9.71	5.25	5.77	8.14	8.060	4.90	7.69	12.48
Well-Matched	10.80	35.38	9.76	13.20	12.07	13.68	8.30	19.68	35.96
Observations	**1,750**	**896**	**533**	**3,607**	**1,326**	**1,479**	**7,589**	**2,236**	**109,227**

Table 3.7

Educational mismatch for female (%)

Educational Mismatch	Black African	Black Caribbean	Chinese	Indian	Other Asian Background	Other Ethnic Group	Other White	Pakistani & Bangladeshi	White British
Over-Educated	80.59	65.16	84.98	80.11	81.9	82.51	89.01	70.75	56.62
Under-Educated	7.48	9.20	5.30	5.16	7.62	7.17	4.23	6.28	11.54
Well-Matched	11.93	25.64	9.72	14.72	10.49	10.32	6.76	22.97	31.83
Observations	**1,978**	**1,478**	**679**	**3,138**	**1,497**	**1,726**	**8,368**	**1,354**	**113,520**

For instance, Black Caribbean males have 9.74% more individuals categorized as well-matched compared to Black Caribbean females. However, there are 10.25% more Black Caribbean females (65.16%) categorized as over-educated compared to Black Caribbean males (54.91%).

Table 3.8

Horizontal and vertical educational mismatch of graduates

Job Mismatch	Males	Females
Horizontal Mismatch	46.50	48.10
Vertical Mismatch	22.30	14.90
Both Horizontal and Vertical Mismatch	9.70	7.30
Well-Matched	41.00	44.30
Observations	**3,011**	**4,221**

Table 3.8 demonstrates the incidence of horizontal and vertical mismatch by gender. Vertical mismatch is also included in this section to gain additional information and fullness for the analysis. At first instance, it is noticeable that the occurrence of horizontal mismatch is much greater than vertical mismatch for both genders. Horizontal mismatch is marginally more likely for females (48.1%) than males (46.5%). However, this is reversed, and the difference is greater in terms of vertical mismatch given that there are 22.3% males mismatched compared to 14.9% for females. Therefore, this implies that females are more prone to be horizontally mismatched whereas males are more likely to be vertically mismatched. Males are slightly more likely to simultaneously be both horizontally and vertically mismatched than females with figures of 9.7% and 7.3% respectively. Finally, females benefit a higher rate of being well-matched than its sex counterpart. Overall, Table 3.8 suggest that female has an advantage of matching their education and occupation than males.

Table 3.9 presents the horizontal and vertical mismatch of graduates by ethnicity and gender. There are recognised characteristic differences among the three Black ethnic groups: Black African, Black Caribbean, and Other Black. However, due to limited sample sizes, these groups are combined and referred to collectively as "Blacks" in the analysis. This grouping is essential to ensure an adequate sample size for more robust estimation.

Including this combined group allows for more detailed insights as opposed to excluding it altogether. In the case of the Black group, all three mismatch categories are predominantly represented by males, underscoring that Black male experience more consistent educational mismatch than females. The disparity is especially striking in horizontal mismatch, where 46% of males and 26.2% of females are in jobs that do not align with their skillset. Additionally, Black females exhibit a significantly higher rate of being well-matched, with a difference of 24.3% compared to males. This indicates that Black females are notably better matched in terms of education and skillset to their occupations compared to their male counterparts.

Similarly, the Other Asian Background and Other Ethnic groups follow a comparable trend to the Black group. These ethnic minority groups also show a higher rate of educational mismatch among males across all three categories. Notably, both the Other Asian Background and Other Ethnic groups exhibit the highest horizontal mismatch rates for males in Table 3.9, with rates of 54.5% and 54.2% respectively. Moreover, Other Asian Background males experience the highest vertical and both mismatch rates, with figures of 31.8% and 18.2% respectively, which are significantly higher than those of other groups. Surprisingly, females from the Other Asian Background group have the highest proportion of being well-matched compared to other ethnic groups, with 60% being well-matched, the same rate as the Bangladeshi and Pakistani group.

In contrast, the outcomes for the Chinese ethnic group are reversed compared to the Black, Other Asian Background, and Other Ethnic groups. Chinese female workers exhibit a higher proportion of horizontal and vertical mismatch compared to their male counterparts, with rates of 50% and 21.4% respectively, whereas males have rates of 38.5% and 15.4% respectively. However, the proportion of both mismatched cases is slightly lower for

Table 3.9

Horizontal and vertical educational mismatch of graduates by gender

Job Mismatch	Blacks		Chinese		Indian		Other Asian Background		Other Ethnic Group		Other White		Pakistani and Bangladeshi		White British	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Horizontal Mismatch	46.00	26.20	38.50	50.00	40.40	37.50	54.50	37.50	54.20	43.80	47.70	51.60	28.90	37.80	46.70	49.10
Vertical Mismatch	31.70	22.40	15.40	21.40	17.50	20.00	31.80	12.50	23.70	12.40	22.10	19.10	24.40	8.90	22.00	14.50
Both Mismatched	9.50	4.70	7.70	7.10	5.30	7.50	18.20	10.00	10.20	2.20	5.80	7.60	6.70	6.70	9.90	7.40
Well - Matched	31.70	56.10	53.80	35.70	47.40	50.00	31.80	60.00	32.20	46.10	36.00	36.90	53.30	60.00	41.20	43.80
Observations	**63**	**107**	**13**	**28**	**57**	**80**	**22**	**40**	**59**	**89**	**86**	**157**	**45**	**45**	**2,666**	**3,675**

females compared to males, although males are much more likely to be well-matched. As for the Indian group, their outcomes are relatively average across all four mismatch categories, with proportions being fairly similar between genders.

Interestingly, the Bangladeshi and Pakistani group shows relatively low levels of educational mismatch overall for both genders. In fact, among males, this group has the lowest rate of horizontal mismatch (28.9%), while females have the lowest rate of vertical mismatch (8.9%).

3.4. Methodology

3.4.1. Earnings equation for vertical mismatch

The earning equation will be assessed to measure the impact to the wage premium for various ethnic groups subject to educational vertical mismatch.

Firstly, let the earnings structure of the following form for individual i:

$$y_i = x_i'\beta + \varepsilon_i \qquad \text{where } i = 1, 2, \dots, \mathrm{N} \tag{3.1}$$

Where y_i represents individual i log of earnings of the, x_i denote the vector of explanatory variables, β are the coefficients representing the impact of each explanatory variable on earnings and ε_i is the random error term.

Now, under the assumption of conditional lognormality of earnings:

$$y_i \mid x_i \sim N(x_i'\beta, \varepsilon_i) \tag{3.2}$$

The equation to estimate the earnings function for various ethnic groups has the form:

$$\ln(y_i) = \beta_0 + \sum_{g=1}^{10} \beta_{1g}\left(EthnicGroup_{ig}\ x\ VerticalUnderEducated_i\right) \tag{3.3}$$
$$+ \sum_{g=1}^{9} \beta_{2g}(EthnicGroup_{ig}\ x\ VerticalWellMatch_i)$$
$$+ \sum_{g=1}^{10} \beta_{3g}(EthnicGroup_{ig}\ x\ VerticalOverEducated_i) + X_i'\beta + \varepsilon_i$$

Where $\ln(y_i)$ represents the dependent variable which denotes the log of hourly earnings which is adjusted for inflation for individual i. The intercept, β_0 represents the constant term that captures the baseline level of the dependent variable when all other variables are zero. The three sums represent interaction terms between different ethnic groups and different educational mismatch statuses. These interaction terms allow the effect of educational mismatch on the dependent variable to vary across different ethnic groups.

The coefficient β_{1g}, β_{2g} and β_{3g} represents the effect of the undereducated, well-matched and overeducated respectively on the dependent variable for ethnic group g. Since the White British group is the base category, there is only 9 ethnic groups for the well-matched term. This is evident from equation (3.3) since overeducated and undereducated term comprise 10 ethnic groups whereas there are only 9 ethnic groups for the well-matched term. X_i represents a set of control variables (such as age, age squared, gender, marital status and a range of job and work-related characteristics such as workplace size, managerial status, number of employees at the workplace, region of workplace etc.[10]). β represents the coefficients for these control variables. Finally, ε_i represents the error term.

[10] The mean and standard deviations of the control variables and other various variable and shown in the appendix.

For each ethnic group, there will be three categories: adequately educated group, over educated group and the under educated group. For ease of interpretation, the adequately educated White British group will be the base category.[11]

The hypothesis is that due to the familiarisation of the economy for the White British group coupled with the potential discrimination to the ethnic minorities, the White British group on average, should earn more than its ethnic minority counterparts. Moreover, overeducated (undereducated) individuals are expected to earn the highest (lowest).

However, educationally well-matched individuals may earn more than the overeducated individuals for several reasons. Firstly, their education level aligns closely with the requirements of their occupation which allows them to effectively utilise their skills, potentially leading to higher productivity, thus earnings. In contrast, difficulty arises for overeducated individuals to utilise their qualifications in their occupations which can lower job satisfaction and productivity. Also, employers may be less inclined to offer the wage align with the individual's level of education as the job does not require the full utilisation of the qualification obtained.

3.4.2. Earnings equation for horizontal mismatch

Horizontal mismatch is when an individual's field of education/expertise does not align with the job. An example is when one occupation requires a different skill or qualification they possess. There are two possible outcomes, either the graduate is horizontally mismatched or horizontally matched. Similar to the vertical mismatch, the wage premium will be examined when the various ethnic groups are either horizontally mismatched or matched. Since the White British group is the ethnic majority, it is logical to use as the base category group.

The hypothesis suggests that among two graduates with identical degrees, the one employed in a position utilising the degree knowledge is anticipated to earn more than the

[11] The White British group is the base category, so it is excluded for the well-matched term. This also ensures to avoid the dummy variable trap.

counterpart in a role where such knowledge is not required. This expectation arises from the assumption that the former individual can leverage their expertise more effectively, thereby contributing more significantly to their respective roles.

To determine whether there is an occurrence of horizontal mismatch, the method of modal fields is adopted from previous papers such as by Chernoff (2012) and Rudakov et al (2019). The modal method measures and compares the graduate's fields of study (i.e. degree subjects – CMBHD01) with respect to their occupation (i.e. industry 2-digit code – INDD07M). The modal method and mismatches are calculated from the statistical measure which reflects the actual distribution of graduates across various fields of education within specific occupations. As a result, graduates are considered matched (mismatched) if they possess the same (different) degree subject as the mode of graduate's employed in the specific occupation.

Therefore, using the equation (3.1) and (3.2) as a foundation, the earnings function to estimate the impact of horizontal mismatch of recent university graduates is as follows:

$$\begin{aligned}\ln(y_i) = \alpha_0 + &\sum_{g=1}^{7} \alpha_{1g}(EthnicGroup_{ig}\ x\ HorizontalWellMatch_i) \\ &+ \sum_{g=1}^{8} \alpha_{2g}(EthnicGroup_{ig}\ x\ HorizontalMismatch_i) + X_i'\beta \\ &+ \varphi HorizontalMismatch + \zeta VerticalMismatch + \varepsilon_i\end{aligned} \tag{3.4}$$

Where $\ln(y_i)$ is the log of hourly earnings which is adjusted for inflation for individual i. This is the constant term in the regression, representing the expected value of $\ln(y_i)$ when all other variables in the model are equal to zero. Horizontal Well Match represents individuals who are well-matched in terms of their education level and job requirements.

The equation includes interaction terms between different ethnic groups and horizontal well-matching. There are 7 ethnic groups in this sum because White British is omitted as the base category, meaning that the effects of the other ethnic groups are compared to White British individuals. α_1 is the coefficient for the interaction between the ethnic group g and horizontal well-matching. This term shows how being horizontally well-matched in terms of education affects the outcome y_i, for each ethnic group relative to the base category (White British).

Horizontal Mismatch refers to individuals whose education field does not match the job they are performing. This term captures the interaction between ethnic group membership and horizontal mismatch (education-job field mismatch). There are 8 ethnic groups in this sum, as it includes all groups, including White British.

X_i represents a set of control variables (such as age, age squared, gender, marital status and a range of job and work-related characteristics such as workplace size, managerial status, number of employees at the workplace, region of workplace etc.[12]). β represents the coefficients for these control variables. φ and ζ represent the coefficients for the variable horizontal mismatch and vertical mismatch respectively. As a result, equation (3.4) base category is the White British group which are horizontally and vertically well-matched.[13]

The horizontal mismatch and vertical mismatch variables from the control variables section provide important information about the general impact of job mismatch (both horizontal and vertical) on the dependent variable for the base category (White British). Therefore, the White British who are horizontally and vertically well-matched are the base group. The coefficients for interaction terms show how much the effect of mismatch (horizontal or vertical) differs for other ethnic groups compared to White British individuals who are well-matched.

[12] The mean and standard deviations of the control variables and other various variable and shown in the appendix.
[13] The base category for Table 3.11 is represented by the White British group which are both horizontally and vertically well-matched.

3.4.3. Determinant of subject choices

Understanding the subject choices chosen and the extent of dissimilarity between ethnic groups is important for policymakers to provide guidance and support to meet the interests of the society. Also, favouring certain subject choices will influence the wage premium and career pathways for graduates.

3.4.3.1. Duncan index

The Duncan Index of Dissimilarity is used to determine the degree choices for various ethnic groups. The purpose of this index is to quantify the degree of certain ethnic minority groups are dissimilar in terms of their choices of university subjects compared with the White British group.

A high Duncan Index value signifies that there is a considerable difference in subject choices between ethnic minority groups and the White British population. This indicates a higher level of dissimilarity, suggesting that members of ethnic minority groups tend to select subjects that are distinct from those chosen by the White British population. Conversely, a low Duncan Index value suggests a lower level of dissimilarity, indicating that subject choices among ethnic minority groups are more aligned with those of the White British population. This implies that members of ethnic minority groups tend to choose subjects that are more similar to those chosen by the White British population.

The estimation equation for the Duncan Index of Dissimilarity can be formulated as follows:

$$DI_{eb} = \frac{1}{2}\sum_{s=1}^{S} \left|\frac{C_{es}}{C_e} - \frac{C_{bs}}{C_b}\right| \quad (3.5)$$

Where DI_{eb} is the Duncan Index of Dissimilarity between ethnic group e and the White British group. C_{es} is the count of individuals from ethnic group e who chose subject s given that C_e is represents the total count of individuals from ethnic group e. C_{bs} is the count of White British individuals who chose subject s, while C_b is the total count of White British individuals. Finally, S represents the total number of subjects (11 total subjects).

3.4.3.2. Multinomial logit model for subject

A multinominal logit model is utilised to investigate the determinants for UK students choosing a certain degree subject in university. This will examine the different ethnic minority choices compared to the White British group. The purpose of this is to understand the propensity of choosing specific degree subjects. In turn, this can provide insight to whether certain ethnic groups are more inclined to choosing subjects that are prone to achieve higher or lower wage premium. For instance, compared to the White British group, does the Indian group have a higher propensity in choosing a degree subject that is linked to STEM subjects.

In this context, the model estimates how individual characteristics represented by X_{ig}, influence the likelihood that an individual i from ethnic group g selects degree subject J.

Following Wooldridge's (2010) standard MNL framework, the model calculates these probabilities, allowing comparison across ethnic groups to assess whether specific groups are more likely to select subjects associated with higher or lower wage premiums (e.g., STEM fields). The modal method is applied to identify the most commonly chosen degree subject for each group, setting the White British group as the base category.

3.4.4. Further investigation of horizontal mismatch

Once the results from the multinominal logit model is obtained, it enables further investigation to the impacts of horizontal mismatch. In particular, there will be two separate estimations. One will only include graduates who studied STEM subjects, while the other will only include graduates who studied non-STEM subjects. For the former estimation, the base category group will be the White British group that have studied STEM subject in university. Similarly, the latter estimation will exclude the White British group, but all individuals included will have studied non-STEM subjects.

3.5. Results and discussion

The regressions will be estimated as a pooled OLS due to the relatively limited number of observations for the ethnic minority groups. The ethnic groups are separated using the interaction terms. Since the Pakistani and Bangladeshi ethnic groups consist relatively small sample size, they are combined into one group, named as “Pakistani and Bangladeshi”. Furthermore, except for Table 3.10, the Black African, Black Caribbean and Other Black group will be grouped together and named as “Blacks”[14] due to small sample sizes. The initial approach involved estimating the regression with distinct categories for Black ethnic groups (Black African, Black Caribbean, and Other Black) to extract as much information as possible. However, limited sample sizes rendered this method ineffective. Consequently, these groups were consolidated under the broader category "Blacks" to enhance the reliability of the analysis.

Table 3.10 presents the vertical mismatch estimation of the earnings equation using pooled Ordinary Least Squares (OLS) regression. The analysis focuses on three categories: Under-educated, Well-matched and Over-educated. These categories allow for comparisons to observe if there are earning premiums associated with the attained education levels. The well-matched White British ethnic group serves as the control group, while the other ethnic minorities represent the treatment groups for comparison.

Similar to Battu and Sloane (2002), the well-matched estimates from Table 2.4 reveal that ethnic minorities generally earn less than White British individuals. Specifically, in the well-matched category, all ethnic minority groups except for the Chinese exhibit an ethnic minority penalty compared to the earnings of the White British group. Interestingly, the Other White group's earnings are significantly lower than those of the White British by 10.3%, suggesting a significant earning penalisation for Caucasian descent who do not belong to specific ethnic groups such as English, Welsh, Scottish, Romani, or Irish.

As previously mentioned, the Chinese group is the only ethnic minority group that demonstrates higher earnings than the White British in the well-matched category. This

[14] Attempted to estimate when the ethnic groups are not combined. The results are not significant, and the standard errors are relatively large due to the limited sample sizes.

finding is supported by data from the Office for National Statistics (2019), which indicates that employees of Chinese ethnicity had a higher median hourly pay than White British employees in 2018, corroborating the results. However, the Chinese groups result is not statistically significant which suggests that the estimate is not substantially different from that of the White British.

On the other hand, the remaining ethnic minority groups categorised as well-matched are statistically significant at the 1% level, indicating that the earnings for each group are significantly different from the White British relative to their coefficient. As a result, the Indian ethnic group has the highest statistically significant earnings compared to other ethnic minority groups, with a 5.5% lower hourly wage than the well-matched White British group. In contrast, the Bangladeshi and Pakistani group has the lowest earnings, with an ethnic minority penalty of 17.8%.

Compared to the White British group, the findings reveal that the well-matched group earns more than the under-educated group but less than the over-educated group, consistent with Rubb (2002), which found that over-educated individuals tend to earn more than their properly educated counterparts. Among White British individuals, there is an average earning penalty of 9.9% for being under-educated for their specific job, whereas there is an average earning premium of 20.6% for individuals who are over-educated compared to those who are well-matched. This pattern extends across ethnic minority groups, where under-educated individuals earn less than the well-matched group, while the over-educated group enjoys a premium, as expected.

Regarding under-educated workers, every ethnic minority group earns less than the White British group and all results are statistically significant at the 1% level (except for Other Black which is significant at the 5% level). Among under-educated ethnic minorities, the Chinese group performs the best, with a 19.7% lower hourly wage than the well-matched White British group. Following closely are the under-educated Black Caribbean and Other Black groups, with earnings lower than the base outcome by 21.2% and 12.8%, respectively. Surprisingly, other than the White British, the Indian group ranks fifth in under-educated earnings, with a penalty of 27.5% compared to the base.

Table 3.10

Regression of wage effects on vertical mismatch and ethnicity

Ethnic Group	Under Educated	Well Matched	Over Educated
Black African	-0.305*** (0.029)	-0.139*** (0.010)	-0.056*** (0.010)
Black Caribbean	-0.212*** (0.034)	-0.077*** (0.010)	0.036** (0.017)
Chinese	-0.197*** (0.063)	0.013 (0.020)	0.195*** (0.015)
Indian	-0.275*** (0.027)	-0.055*** (0.007)	0.118*** (0.007)
Other Asian Background	-0.367*** (0.031)	-0.154*** (0.011)	-0.045*** (0.012)
Other Black	-0.218** (0.097)	-0.043 (0.027)	0.055** (0.033)
Other Ethnic Group	-0.308*** (0.028)	-0.128*** (0.011)	0.022** (0.010)
Other White	-0.236*** (0.020)	-0.103*** (0.005)	0.032*** (0.005)
Pakistani and Bangladeshi	-0.346*** (0.030)	-0.178*** (0.010)	-0.001 (0.010)
White British	-0.099*** (0.003)	- -	0.206*** (0.002)
Control Variables			
Age	0.045*** (0.000)		
Age Squared (x1000)	0.472*** (0.170)		

Female	-0.094*** (0.000)
Tenure	0.018*** (0.000)
Tenure Squared	0.000*** (0.000)
Constant	**1.591*** -0.015**
Observations	**260,010**
R-Squared	**0.3957**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to two (three) decimal places. The base category is the "Vertically – Matched White British" group.

Conversely, the Other Asian Background and the Bangladeshi and Pakistani group have the lowest under-educated earnings, with an hourly wage differential of 36.7% and 34.6%, respectively, compared to the base.

In the overeducated category, the Chinese ethnic group continues to outperform its ethnic minority counterparts with a wage premium of 19.5% compared to the base. Similarly, the White majority enjoys a wage premium benefit for overeducated workers, obtaining a 20.6% higher hourly wage compared to its well-matched White British counterparts. Interestingly, unlike the undereducated category, the overeducated Indian group ranks second in terms of earnings among ethnic minorities, with a wage premium of 11.8% compared to the base.

Regarding the three overeducated, right level of education, or undereducated (ORU) specification, it is evident that the Chinese group's earnings exceed those of the Indian group in every category, corroborating with data from the Office for National Statistics (2020) which illustrates that the Chinese ethnic group has a greater median hourly pay than the Indian group.

Conversely, there are three overeducated ethnic minority groups that earn lower than the well-matched White British group. The overeducated Black African and Other Asian

Background groups have an hourly wage 5.6% and 4.5% respectively lower than the base groups. Additionally, the Bangladeshi and Pakistani group has an earnings penalty of -0.1% compared to the base, although the estimate is not statistically significant, likely due to the relatively homogenous coefficient compared to White British earnings.

Moreover, Groot and Brink (2002) found that there is only little systematic variation in the rate of return to overeducation and undereducation. Our estimation aims to determine the diverse propensity to variation of the rate of return in dependence for each ethnic group of interest. Consistent with Groot and Brink (2002), the findings suggest that overeducated workers earn a premium, whereas undereducated workers suffer from a wage penalization compared to well-matched workers. They concluded that the 'true' rate of return to a year of overeducation is 2.6%, while the rate of return to a year of undereducation is -4.9%, indicating that, on average, workers who are under-educated experience a greater magnitude of penalty compared to the gain in premium for workers who are over-educated.

From Table 3.10, it is clear that undereducated workers face a substantially larger penalty than overeducated workers receive as a premium. This finding is consistent with the findings of Groot and Brink (2002). The magnitude of the penalty (or premium) is determined by the difference in the hourly wage rate between the well-matched and undereducated (or overeducated) categories. For example, focusing on the Indian ethnic group, undereducated workers experience a wage penalty of 22%, whereas overeducated workers earn 17.3% more than their well-matched counterparts. Apart from the White British group, only two ethnic groups—Other White (0.2%) and Bangladeshi and Pakistani (0.9%)—show a magnitude where the premium for overeducated workers is greater than the penalty for undereducated workers relative to the well-matched group.

In summary, the Pooled Ordinary Least Squares (POLS) estimate shown on Table 3.10 aligns with previous studies, showing that when comparing workers with the same attained level of education, overeducated workers earn less, and undereducated workers earn more than their well-matched counterparts. As expected, female workers earn less than males, and the overeducated penalty is more pronounced for females.

Regarding ethnicity, the general hypothesis that White British workers earn more on average is supported. The Chinese ethnic group stands out as the only group with a higher hourly wage rate than the White British. The Indian group earns slightly less than the ethnic majority, and as anticipated, the Bangladeshi and Pakistani group faces the largest ethnic minority penalty, consistent with this research findings.

Thus, similar to Battu and Sloane (2002), the results indicate the presence of vertical mismatch between educational qualifications and occupational attainment. The White British group consistently earns significantly more than all ethnic minorities across every ORU specification. When compared to the base outcome category (well-matched), undereducated workers experience a wage penalty, while overeducated workers enjoy a wage premium. Furthermore, strong evidence of heterogeneity between ethnic minorities is observed, with wage differentials present in all three ORU categories. Consistent with previous research, the Chinese and Indian groups tend to earn higher wages, whereas the Black African, Other Ethnic, and Bangladeshi and Pakistani workers experience the greatest wage penalties in each category.

Table 3.11 presents the estimation of the logarithm of gross hourly earnings of graduates, with a focus on assessing the impact of various ethnic groups on earnings magnitude. The ethnic minorities are compared to the base category control group - the White British who are horizontally well matched. The two control variables, horizontal mismatch and vertical mismatch examines the effect of educational mismatch on the hourly earnings of the base category. The coefficient for the horizontal (vertical) mismatch variable (without interaction) captures the overall effect of being horizontally (vertically) mismatched for the base category (i.e. White British). It shows how being horizontally (vertically) mismatched affects the outcome for the reference group (White British), controlling for other variables.[15]

The field of occupation is not included as a control in the regression analysis to avoid conflating the effects of potential discriminatory practices with pure wage effects. Ethnic minorities often experience occupational clustering, where they are disproportionately represented in certain lower-paying or less prestigious occupations. Controlling for occupation could mask or obscure the effects of discrimination or structural inequalities

[15] The interaction terms allow the model to capture heterogeneous effects of horizontal (vertical) mismatch across ethnic groups. The separate horizontal mismatch variable represents the base effect for the reference group (White British), and the interaction terms show deviations from that baseline for ethnic minority groups.

Table 3.11

Graduate wage effects by horizontal mismatch and ethnicity

Ethnic Group	Horizontally Matched	Horizontally Mismatched
Blacks	-0.100*** (0.040)	-0.074 (0.069)
Chinese	0.127 (0.090)	-0.010 (0.133)
Indian	0.023 (0.046)	0.033 (0.073)
Other Asian Background	-0.035 (0.076)	-0.099 (0.112)
Other Ethnic Group	-0.004 (0.049)	-0.119* (0.071)
Other White	-0.006 (0.039)	-0.012 (0.055)
Pakistani and Bangladeshi	-0.009 (0.054)	-0.087 (0.091)
White British	- -	-0.080 (0.010)
Control Variables		
Age	0.058*** (0.004)	
Age Square	-0.062*** (0.005)	
Female	-0.118*** (0.010)	
Tenure	-0.017** (0.009)	
Tenure Square	0.004***	

	(0.001)
Horizontal Mismatch	-0.030*** (0.010)
Vertical Mismatch	-0.108*** (0.013)
Constant	**1.563*** (0.089)**
Observations	**6,774**
R-Squared	**0.3484**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places. The base category is the "Horizontally – Matched White British" group.

that may lead ethnic minorities to be horizontally mismatched in the first place. By excluding occupation as a control variable, the analysis aims to capture the full impact of horizontal mismatch on wages across ethnic groups, without adjusting for the unequal distribution of occupations that could itself be a result of discriminatory labour market practices.

It is evident that both educational mismatch variables have a negative impact on hourly earnings, as predicted and observed previously. The results indicate that vertical mismatch has a significantly greater negative effect than horizontal mismatch, reducing hourly earnings by 10.8% and 3% respectively. Both coefficients are statistically significant at the 1% level, implying a substantial impact on hourly earnings. These findings are supported by Mahuteau, et al. (2014), who report that horizontal mismatch alone does not substantially reduce hourly wages, whereas vertical mismatch alone can significantly lower hourly wages.

In fact, they found that horizontal mismatch alone does not lower hourly wages for men, suggesting that vertical matching is more crucial for graduates than horizontal matching. As hypothesised, the combination of horizontal mismatch and vertical mismatch contributes to the largest wage penalty.

As observed from the nominal ethnic minority category (non-horizontal), in most cases, they incur lower hourly earnings than their White British counterpart. Although, it should be mentioned that the ethnic minority groups are compared to the White British group which are well-matched since the analysis includes the horizontal mismatch and vertical mismatch explanatory variables. Therefore, the results for the ethnic minority will be negatively biased since the ethnic minority is horizontally well matched but includes samples which are vertically mismatched. Unsurprisingly, the Blacks[16] ethnic group possess the greatest wage penalty of -10 % in comparison to the control group. This is the only result which is statistically significant which is likely due to a lack of sample size. Even though this is a drawback from the decrease in statistical power, the coefficients of the overall result are still insightful and align with the expected result.

Moreover, the Bangladeshi and Pakistani, Other Asian Background, and Other White ethnic groups also experience lower hourly earnings than the White British group by 0.9%, 3.5%, and 0.6% respectively. While it's expected that these groups would earn less per hour than the White British, it's surprising that the Other Ethnic group shows the same hourly earnings as the White British group. However, as previously mentioned, this outcome may be influenced by a lack of sample size, necessitating further analysis.

Conversely, based on the coefficients in Table 3.11, the Chinese and Indian ethnic groups perform better than the White British group. This result aligns with expectations and is consistent with previous research and the earlier findings from Table 3.10. The analysis suggests that compared to the White British group, the Chinese and Indian group have 12.7% and 2.3% higher gross hourly earnings, respectively. The magnitude of these coefficients implies that the Chinese group earns significantly more than the Indian, White British, and other ethnic minority groups.

The 'Horizontal' category within ethnic minorities comprises groups experiencing horizontal mismatches. The majority of the horizontally mismatched group tend to earn lower hourly wages compared to their horizontally matched counterparts. However, exceptions to this trend are observed within the Indian and Black demographic. The Other Asian and Other

[16] Black African, Black Caribbean and Other Black are grouped together and named as Blacks for Table (3.11). This is due to low sample sizes which cause a low statistical power which cause difficulty to draw reliable and meaningful conclusions from the data.

Ethnic categories face the greatest penalties, with disparities of 11.9% and 9.9%, respectively, for being categorised as ethnic minorities and experiencing horizontal mismatches. In the 'horizontal' category, the Other Ethnic group is the only result which is statistically significant. This suggests the earlier conclusion about the similarity in hourly earnings between the nominal Other Ethnic group and the White British group is likely inaccurate due to the limited sample size.

Consistent with prior findings, both the Black, as well as the Bangladeshi and Pakistani groups, experience significant penalties of 7.4% and 8.7%, respectively, compared to the White British group. Interestingly, Blacks perform better when they are horizontally mismatched, as indicated in Table 3.11. This anomaly is likely due to small sample sizes, which reduce the ability to detect an effect. However, since all coefficients except for the Indian group are negative, there is evidence suggesting that groups characterized as horizontally mismatched and belonging to an ethnic minority face penalties in their hourly earnings.

Additionally, explanatory variables, horizontal and vertical mismatch, are introduced to examine the impact of educational mismatch among the White British group. Consistent with Mahuteau et al. (2014), the findings reveal that vertically mismatched individuals face a significantly higher penalty compared to horizontally mismatched individuals, by 10.8% and 3%, respectively. It's notable that Mahuteau et al. found that horizontal mismatch alone doesn't reduce hourly earnings for men, whereas vertical mismatch alone does. However, as anticipated and in line with previous research, individuals experiencing both horizontal and vertical mismatch generally incur the greatest penalty in hourly earnings.

Overall, the other independent variables align with typical characteristics found in previous studies. For instance, female graduates in the analysis earn substantially less than their male counterparts, by 11.8%. With the exception of the variable "Number of Dependent Children," most independent variables are statistically significant at the 1% and 5% levels. The variable "Degree Class" indicates that graduates achieving a lower degree class than first-class honours earn considerably less. The earning penalty tends to increase as the degree class obtained is lower, except for the "Pass" classification, although this coefficient lacks statistical significance, thus failing to reject the null hypothesis of no effect on hourly earnings.

Non-managerial positions tend to incur substantial earnings penalties, with figures reaching 19.7% and 29.8%, particularly notable for graduates who are not managers or supervisors. Additional labour market factors, such as firm size, also influence graduate hourly earnings. There's a positive correlation between firm size, measured by the number of employees, and individual hourly earnings. For example, graduates working in firms with over 500 employees earn, on average, 29% more than those in firms with 1-10 employees.

Workplace location is another crucial determinant of graduate hourly earnings, aligning with standard findings from previous studies. Graduates in Central London receive the highest hourly earnings, followed by those in the Rest of Yorkshire and Humberside. Inner and Outer London are associated with the next highest earnings. Conversely, the Merseyside area reports the lowest hourly earnings, with salaries lower than the region base category of Tyne and Wear by 1%.

Similar to the previous regression, Table 3.12 presents estimations for hourly earnings accounting for effects from horizontal mismatch. However, in this regression, horizontally mismatched ethnic minorities have been categorised into two distinct groups. The first category, labelled as the "strong positive horizontal," comprises the Chinese and Indian ethnic groups, which are relatively high earners despite being horizontally mismatched.

The aim of this analysis is to determine if grouping such ethnic minorities provides additional insights and robustness to the previous findings from Table 3.11. The second category, the "strong negative horizontal" group, encompasses ethnic minorities experiencing a significant penalty from horizontal mismatch. It's important to note that, since the Other White group possesses similar characteristics to the White British, this group has been maintained separately from the strong negative/positive horizontal groups.

From Table 3.12, the strong positive horizontal group indicates slightly higher hourly earnings compared to the well-matched White British group, by 2.3%. However, this coefficient lacks statistical significance. This result suggests two potential outcomes. Firstly, there may not be a sufficient number of observations, leading to low statistical power.

Table 3.12

Graduate wage effects with strong negative/positive by horizontal mismatch and ethnicity

Ethnic Groups	Horizontally Matched	Horizontally Mismatched
Blacks	-0.093*** (0.035)	
Chinese	0.111 (0.073)	
Indian	0.027 (0.044)	
Bangladeshi and Pakistani	-0.006 (0.046)	
Other Asian Background	-0.037 (0.059)	
Other Ethnic Group	-0.016 (0.041)	
Other White	-0.006 (0.039)	
Well – Matched White British	- -	
Other White Horizontal		-0.012 (0.055)
Strong Negative Horizontal		-0.095** (0.041)
Strong Positive Horizontal		0.023 (0.064)
Control Variables		
Age	0.058*** (0.004)	
Age Square	-0.062***	

	(0.005)
Female	-0.118*** (0.010)
Tenure	-0.017** (0.009)
Tenure Square	0.004*** (0.001)
Horizontal Mismatch	-0.030*** (0.010)
Vertical Mismatch	-0.108*** (0.013)
Constant	**-0.156*** (0.089)**
Observations	**6,774**
R-Squared	**0.3995**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places. The base category is the "Well – Matched White British" group.

A limited sample size could result in a large standard deviation, reducing result accuracy due to poor representation of the population. Secondly, it's plausible that the true effect of the strong horizontal group doesn't significantly deviate from the White British group's earnings. Previous studies (Mahuteau et al. 2014) suggest that the former outcome is more likely, given that the Chinese and Indian groups tend to experience statistically significant higher earnings than the White British group despite being characterised as horizontally mismatched.

The strong negative horizontal group exhibits lower hourly earnings of 9.5% compared to the White British group. This result is statistically significant at the 5% confidence level, suggesting that members of the horizontally mismatched ethnic minority group are highly likely to experience lower hourly earnings compared to the White British group. Clustering the low-earning ethnic minorities together increases the sample size of the dummy variable, leading to a statistically significant result.

In contrast, the strong positive horizontal dummy variable appears less effective. Despite the positive coefficient, the statistical power remains weak. This is likely due to the limited sample size resulting from aggregating only two horizontally mismatched ethnic minorities.

Table 3.13 presents the Duncan Index, also known as the Dissimilarity Index, which indicates the segregation between ethnic groups based on the subjects chosen for degrees. In the analysis, the White British group is used as the reference group and compared it with each ethnic minority group. The index reflects how different the distribution of degree subjects is between each ethnic group and the White British group.

The results reveal that the Chinese group has the highest Dissimilarity Index compared to the White British group, calculated at 0.379. This suggests that nearly 38% of Chinese workers would need to choose a different degree subject to achieve a distribution similar to that chosen by the White British. Similarly, the Blacks, Indian, Bangladeshi and Pakistani, and Other Asian groups show comparable Dissimilarity Index values, ranging from 32.2% to 36.8%.

Table 3.13

Effect of subject degree using the Duncan index

Ethnicity	**Blacks**	**Chinese**	**Indian**	**Other Asian**	**Other Ethnic Group**	**Other White**	**Bangladeshi and Pakistani**
Duncan Index	0.3489	0.3785	0.3456	0.3220	0.1850	0.1923	0.3683
Observations	**6,511**	**6,382**	**6,478**	**6,403**	**6,489**	**6,584**	**6,431**

Note: The base category group is the White British group. Therefore, the index shows the dissimilarity with respect to the White British group.

In contrast, the Other Ethnic and Other White groups exhibit significantly lower Duncan Index values, at 0.185 and 0.192 respectively. This indicates that only around 18.5% to 19.2% of individuals from these groups would need to choose a different degree subject to achieve a distribution similar to that of the White British. The similarity in subject degree distribution for the Other White group may be attributed to individuals from European countries sharing similar values and occupational mindsets with the White British.

Conversely, data from HESA (2014 – 2019) shows that the Chinese group predominantly chooses STEM (Science, Technology, Engineering, and Mathematics) subjects such as Engineering, Medicine, and related fields. While a significant proportion of the White British group also chooses STEM subjects, they are more likely to opt for non-STEM subjects than the Chinese group, resulting in a substantial distribution disparity.

The Duncan Index of Dissimilarity highlights disparities in degree subject choices among ethnic minority groups when compared to the White British group. To further analyse the differences, a multinomial logistic regression is estimated using Table 3.14 to examine the likelihood of each ethnic group selecting specific degree subjects.

Initially, the analysis covered 18 subjects. However, due to limitations in sample size, similar subjects were grouped together, resulting in 11 categories. This adjustment ensures statistical robustness, adhering to Schwab's (2002) recommendation of a minimum of 10 observations per independent variable.

The reported marginal effects in Table 3.14 are all relative to a baseline individual: a male of White British ethnicity, with a first-class undergraduate degree, educationally well-matched, in a managerial position, and studying Humanities and Languages. For example, a Black individual is 4.9% less likely to choose Biological Studies but 2.7% more likely to opt for Mathematical Sciences and Computing compared to this baseline, as shown in Table 3.14 (a).

Initial observation reveals the multinomial logit analysis indicates a general inclination among ethnic minority groups towards choosing STEM (Science, Technology, Engineering, and Mathematics) subjects over the White British group. This trend is evident in Table 3.14 (a) – (d), where the marginal coefficients for ethnic minorities in STEM subjects tend to be positive, indicating a higher likelihood of selecting these subjects compared to the ethnic majority.

Notably, the Chinese group exhibits the highest propensity for choosing STEM subjects among ethnic minorities. For instance, they demonstrate the highest marginal probabilities for subjects such as Medicine and Medical related subjects, Computing, and Engineering and Technology, with coefficients of 10.4% and 7.8% respectively. The only exception is Biological Sciences, where the Chinese group ranks second, with a marginal probability of 11.2% for Mathematical Sciences and Computing.

Similarly, the Indian group also shows a relatively strong inclination towards STEM subjects. This alignment with STEM subjects is consistent with previous findings, as both the Chinese and Indian groups are known for their higher earnings and tendency to work in STEM-related occupations post-graduation.

Interestingly, the Other Asian Background group also exhibits a notable likelihood of studying STEM subjects. For instance, they have the second-highest marginal probabilities for Medicine and Medical related subjects, as well as for Biological Sciences, with coefficients of 8.3% and 6.1% respectively. Moreover, the Other Asian Background group stands out as the most likely ethnic group to choose Mathematical Sciences and Computing, with a 12.5% higher probability compared to their White British counterparts.

In contrast, the Other White ethnic group shows a distinct pattern, exhibiting less inclination towards selecting STEM subjects compared to other ethnic minority groups. Instead, its subject selection tendencies closely mirror those of the White British cohort. For instance, across STEM subjects, the marginal probability differences between the Other White group and the White British group range between -0.9% and 0.9%. This suggests minimal divergence in subject preferences between these two groups.

Furthermore, the Other White group stands out as the only ethnic minority group more inclined to opt for non-STEM subjects compared to both the White British group and other ethnic minorities. These findings are consistent with those from the Duncan Index of Dissimilarity in Table 3.13, indicating a similarity in subject selection behaviour between the Other White and White British groups. This similarity may be attributed to their shared White ethnic background and potentially similar socio-cultural influences.

Table 3.14 (a)

Multinominal logit – Subjects

Subjects	Medicine and Medical Related Subjects		Biological Sciences		Agricultural and Physical Sciences		Mathematical Sciences and Computing	
Variables	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**
Age	-0.044	-0.003	0.019	0.000	0.084**	0.003	0.028	-0.001
Age Square	0.049	0.004	-0.018	0.001	-0.090*	-0.003	-0.028	0.001
Tenure	0.007	0.000	0.080	0.004	-0.052	-0.003	0.004	0.000
Tenure Square	0.006	0.000	0.000	0.000	0.005	0.000	0.004	0.000
Blacks	1.588***	0.046	-0.587	-0.049	-0.082	-0.026	0.903**	0.027
Chinese	1.934***	0.104	0.380	0.016	1.535**	0.074	2.035***	0.112
Indian	2.446***	0.071	1.359***	0.040	0.528	-0.009	1.896***	0.072
Bangladeshi and Pakistani	1.385***	0.023	1.579***	0.057	0.282	-0.018	0.848	0.008
Other Asian Background	1.330**	0.083	0.940	0.061	1.059**	0.063	1.957***	0.125
Other Ethnic Group	0.503	0.007	0.737**	0.033	0.127	-0.002	0.676**	0.028

Continuation of Table 3.14 (b)

Variables	Coeff.	Marg E.	Coeff.	Marg E.	Coeff.	Marg E.	Coeff.	Marg E.
Other White	-0.269	-0.007	-0.275	-0.009	-0.876**	-0.047	0.048	0.009
White British	-	-	-	-	-	-	-	-
Female	0.165	0.021	-0.431***	0.001	-1.133***	-0.041	-1.269***	-0.056
Upper Second 2:1	-0.666***	-0.029	-0.469***	-0.016	-0.394**	-0.011	-0.684***	-0.036
Lower Second 2:2	-0.717***	-0.033	-0.275	-0.009	-0.147	0.000	-0.610***	-0.035
Third	-0.298	-0.029	0.316	0.005	0.484**	0.017	0.702***	0.050
Pass	1.910***	0.127	0.153	-0.029	0.622**	-0.006	1.109***	0.038
Horizontal Mismatch	-0.853***	-0.016	1.223***	0.076	1.302***	0.084	0.468	0.052
Vertical Mismatch	-0.612***	-0.026	0.093	-0.004	0.168	-0.002	0.146	-0.004
Foreman or supervisor	0.863***	0.054	0.461	0.023	0.360**	0.016	-0.208	-0.017
Not manager or supervisor	-0.326**	-0.008	-0.121	-0.003	0.038	0.006	-0.050	0.002
Constant	**-0.649**		**-3.119**		**-3.708*** **		**-1.727*** **	

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The excluded categories: Male, First Class Honours, Manager or Supervisor. The marginal effect for a continuous variable is calculated for a one-unit increase. For interaction variables it represents an average person with that particular characteristic relative to the base characteristic. The base category is the “White British” and the “Humanities and Languages” group.

Continuation of Table 3.14 (c)

Subjects	Engineering and Technology		Architecture and Related Subjects		Social Sciences		Business and Financial Studies	
Variables	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**
Age	0.147***	0.006	0.022	0.000	0.017	-0.001	0.096***	0.008
Age Square	-0.154***	-0.006	-0.077	0.000	-0.016	0.003	-0.142***	-0.013
Tenure	-0.050	-0.003	-0.002	0.000	-0.075	-0.011	0.114	0.015
Tenure Square	0.011	0.001	-0.009	0.000	0.003	0.000	-0.009	-0.001
Blacks	1.250***	0.040	1.008	0.004	1.190***	0.074	1.788***	0.141
Chinese	2.016***	0.078	2.618**	0.011	0.902	0.212	1.915***	0.238
Indian	1.734***	0.050	-15.67	-0.079	0.788**	-0.023	1.951***	0.135
Bangladeshi and Pakistani	1.662***	0.051	-15.525	-0.078	1.463***	0.068	2.201***	0.160
Other Asian Background	0.969	0.037	-15.913	-0.078	1.168***	0.264	1.474***	0.202
Other Ethnic Group	0.792**	0.029	-16.297	-0.080	0.812***	0.054	0.737***	0.043

Continuation of Table 3.14 (d)

Variables	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**
Other White	0.193	0.015	1.200**	0.006	-0.038	0.007	0.348**	0.052
White British	-	-	-	-	-	-	-	-
Female	-2.385***	-0.097	-0.436	0.000	0.115	0.048	-0.661***	-0.049
Upper Second 2:1	-0.595***	-0.024	-0.573	-0.002	0.132	0.026	0.186	0.034
Lower Second 2:2	-0.511***	-0.023	-0.276	-0.001	0.266**	0.038	0.129	0.019
Third	-0.002	-0.019	-16.766	-0.007	0.322	0.007	0.460**	0.022
Pass	0.888***	0.009	-16.284	-0.007	0.571**	-0.037	1.213***	0.046
Horizontal Mismatch	0.697***	0.054	16.872	0.011	-3.091***	-0.266	-1.394***	-0.114
Vertical Mismatch	1.025***	0.055	1.010***	0.005	-0.029	-0.022	0.670***	0.077
Foreman or supervisor	0.267	0.011	0.445	0.002	-0.113	-0.017	-0.422***	-0.053
Not manager or supervisor	0.076	0.009	0.060	0.001	-0.211**	-0.009	-0.353***	-0.029
Constant	**-4.682*** **		**-19.644**		**-0.251**		**-2.027*** **	

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The excluded categories: Male, First Class Honours, Manager or Supervisor. The marginal effect for a continuous variable is calculated for a one-unit increase. For interaction variables it represents an average person with that particular characteristic relative to the base characteristic. The base category is the “White British” and the “Humanities and Languages” group.

Continuation of Table 3.14 (e)

Subjects	Librarianship and Information Studies		Humanities and Languages		Education	
Variables	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**	**Coeff.**	**Marg E.**
Age	-0.323**	-0.001	-	-0.009	-0.018	-0.002
Age Square	0.419**	0.001	-	0.010	0.038	0.003
Tenure	0.417	0.001	-	-0.001	-0.025	-0.001
Tenure Square	-0.032	0.000	-	0.000	0.006	0
Blacks	-15.227	-0.047	-	-0.139	-0.794	-0.071
Chinese	-15.701	-0.048	-	0.064	-16.740	-0.863
Indian	-15.128	-0.048	-	-0.224	1.111**	0.017
Bangladeshi and Pakistani	-14.8	-0.046	-	-0.219	0.738	-0.005
Other Asian Background	-15.243	-0.045	-	0.134	-16.513	-0.847
Other Ethnic Group	-15.698	-0.048	-	-0.068	0.449	0.005

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The excluded categories: Male, First Class Honours, Manager or Supervisor. The marginal effect for a continuous variable is calculated for a one-unit increase. For interaction variables it represents an average person with that particular characteristic relative to the base characteristic. The base category is the “White British” and the “Humanities and Languages” group.

Continuation of Table 3.14 (f)

Variables	Coeff.	Marg E.	Coeff.	Marg E.	Coeff.	Marg E.
Other White	0.773	0.003	-	0.034	-1.285**	-0.062
White British	-	-	-	-	-	-
Female	0.600	0.003	-	0.120	0.895***	0.051
Upper Second 2:1	0.500	0.001	-	0.047	0.141	0.010
Lower Second 2:2	1.064	0.003	-	0.023	0.351	0.018
Third	1.035	0.002	-	-0.073	0.686**	0.023
Pass	-15.664	-0.001	-	-0.175	1.448***	0.035
Horizontal Mismatch	16.542	0.006	-	0.161	-1.466***	-0.048
Vertical Mismatch	-0.527	-0.002	-	-0.044	-0.753***	-0.033
Foreman or supervisor	-0.605	-0.002	-	-0.020	0.078	0.002
Not manager or supervisor	0.123	0.001	-	0.034	-0.226**	-0.005
Constant	**-17.632**		**-**		**-2.032****	

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The excluded categories: Male, First Class Honours, Manager or Supervisor. The marginal effect for a continuous variable is calculated for a one-unit increase. For interaction variables it represents an average person with that particular characteristic relative to the base characteristic. The base category is the “White British” and the “Humanities and Languages” group.

As mentioned earlier, both the Black and Bangladeshi/Pakistani ethnic groups tend to earn less on average. However, the data from the Table indicates that these groups are more inclined to choose STEM subjects than their White British counterparts. In particular, the Bangladeshi and Pakistani group exhibit the second-highest likelihood of selecting Biological Sciences and Engineering and Technology, with marginal coefficients of 5.7% and 5.1%, respectively. Similarly, the Black group falls within the median range compared to other ethnic minority groups. However, overall, the Black group is more likely than the White British group to choose each of the STEM subjects except for Biological Sciences.

Conversely, there is evidence suggesting that ethnic minorities, are less inclined to choose non-STEM subjects compared to the base group. This trend is especially pronounced in subjects such as Agricultural and Physical Sciences, Librarianship and Information Studies, Humanities and Languages, and Education Studies. For instance, all ethnic minority groups, except the Other White group, show a decreased likelihood of choosing Librarianship and Information Studies compared to the White British group.

Interestingly, the results show an unexpected pattern for the Social Sciences and Business and Financial Studies subjects. Unlike other non-STEM subjects, ethnic minority groups are generally more likely to choose these two subjects compared to the White British group. This anomaly could be attributed to the classification of these subjects. Definitions of STEM subjects can vary between different organisations and countries, leading to complexities in categorisation. For instance, while the UK's Joint Academic Coding System (JACS) does not classify Social Studies or Business and Administrative Studies as STEM subjects, organisations in the United States, such as the National Science Foundation and Department of Labour, often include Social Sciences and Economics in STEM classifications (Gonzalez, 2012).

Consequently, as expected, ethnic minorities show a stronger tendency to choose Social Sciences and Business and Financial Studies. These subjects exhibit the highest marginal probabilities among the options, particularly for the Chinese and Other Asian Background groups. Specifically, compared to the White British group, the Chinese and Other Asian Background groups are 21.2% and 26.4% more inclined to select Social Sciences, respectively, and 26.8% and 20.2% more inclined to choose Business and Financial Studies, respectively.

Aside from ethnic background, other factors such as gender significantly influence an individual's likelihood of choosing a particular subject. Table 3.14 reveals gender biases since certain subjects are disproportionately selected based on gender. For instance, females statistically show a higher likelihood of selecting Medicine and Medical related subjects, Social Sciences, Humanities and Languages, as well as Education. The dominance of females is particularly pronounced in subjects like Humanities and Languages, where the marginal probability is 12.0% greater than that for males. This finding aligns with previous studies (HEPI, 2014), which have also highlighted these subjects as significantly more popular among females.

Results from Table 3.14 reveal that males are more inclined to choose Agricultural and Physical Sciences, Mathematical Sciences and Computing, Engineering and Technology, as well as Business and Financial Studies. This aligns with findings from HEIPR (2014), which analysed the HEFCE dataset. However, there are subjects where gender differences are less pronounced, such as Biological Sciences, Architecture and related subjects, and Librarianship and Information Studies. Overall, Table 3.14 suggests a gender imbalance in STEM subject choices, with males being more prominent in opting for STEM degrees. This concurs with HESA's (2019) findings, indicating that females are underrepresented in STEM degree subjects, with only 42% of female undergraduates enrolling in science subjects compared to 52% of males in 2017/18.

Additionally, evidence suggests that individuals opting for STEM subjects are more likely to attain a first-class degree, suggesting a reverse causality effect where the choice of STEM significantly impacts the degree class achieved. This pattern is particularly prominent in fields like Medicine, Biological Sciences, and Engineering. Most non-STEM subjects, barring Architecture and related fields, tend to exhibit a positive marginal probability for degree classes below first-class. However, the evidence for Architecture and related subjects is inconclusive. Research from the Educational Testing Service in the United States suggests that students pursuing degrees in Physics, Mathematical Sciences, and Philosophy demonstrate the highest IQ levels, hinting at the possibility of an individual's IQ influencing their choice between STEM and non-STEM subjects on a broader scale.

Table 3.15

Graduate wage effects with strong negative/positive by horizontal mismatch and ethnicity – STEM graduates only

Ethnic Group	Horizontally Matched	Horizontally Mismatched
Blacks	-0.144 (0.091)	
Chinese	0.033 (0.125)	
Indian	-0.067 (0.082)	
Other Asian Background	0.006 (0.125)	
Other Ethnic Group	-0.077 (0.095)	
Other White	0.035 (0.114)	
Pakistani and Bangladeshi	-0.017 (0.107)	
Well – Matched White British		- -
Other White Horizontal		0.017 (0.136)
Strong Positive Horizontal		0.185* (0.104)
Strong Negative Horizontal		-0.100 (0.090)
Control Variables		
Age	0.065*** (0.009)	

Age Square	-0.066*** (0.010)
Female	-0.133*** (0.022)
Tenure	-0.053*** (0.019)
Tenure Square	0.006*** (0.002)
Horizontal Mismatch	-0.135*** (0.024)
Vertical Mismatch	-0.124*** (0.026)
Constant	**-0.164*** (0.193)**
Observations	**1,537**
R-Squared	**0.3996**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places. The base category is the "Horizontally – Matched White British" group.

Table 3.15 presents the estimation results of the hourly earnings equation with the inclusion of ethnic minority horizontal mismatch dummy variables, focusing exclusively on graduates who studied STEM subjects. Similar to Table 3.12, the regression includes the strong positive and negative groups. This analysis determines whether limiting the observations to STEM graduates statistically affects ethnic minorities hourly earnings compared to the baseline estimation of Table 3.12. The hypothesis suggests that by considering only STEM graduates, ethnic minority coefficients will decrease on average due to the relatively greater premium enjoyed by the White British group. This idea stems from the assumption that in a scenario where all graduates study STEM subjects, the White British group will proportionally benefit from a greater premium compared to ethnic minorities in the UK job market.

The results from Table 3.15 indicate that including only STEM graduates reduces the coefficient for nominal ethnic minorities overall. This supports the hypothesis, as most ethnic minorities hourly earnings decreased, particularly among relatively high-earning ethnic groups. For example, the Chinese and Indian groups experienced an earnings penalty of 8.6% and 10.3%, respectively, compared to the results from Table 3.12.

However, there are occasions where including only STEM graduates increased the hourly earnings coefficient for certain ethnic minorities compared to the White British group. For instance, the Other Asian Background and Other White coefficients increased by 5.2% and 4.9%, respectively, compared to Table 3.12. These are anomalies as they represent the only ethnic minorities where hourly earnings increased compared to the White British group. Overall, this suggests that the hourly earnings penalty gap between nominal ethnic minorities reduces when only STEM graduates are considered, as opposed to including all graduates. A possible explanation may be that employers perceive STEM graduates as reputable regardless of the individual's ethnic background.

The strong positive horizontal result in Table 3.15 shows a substantially greater hourly earnings coefficient compared to the result in Table 3.12. In the regression exclusively including STEM graduates, the strong positive horizontal indicates higher hourly earnings by 15.6% compared to Table 3.12. This suggests that the nominal hourly earnings for ethnic minorities are underestimated, given that the coefficients for Chinese and Indian groups in Table 3.15 are 3.3% and -6.7%, respectively. With the strong positive horizontal presenting a coefficient of 18.5% and statistical significance at the 10% level, this value is likely closer to the true representation of hourly earnings for ethnic minority high earners. Therefore, this reinforces the notion that incorporating strong horizontal dummy variables provides further insight and robustness beyond the nominal results.

Table 3.16 presents the hourly earnings estimation exclusively for individuals who studied non-STEM subjects during higher education. Initially, it's evident that the Strong Negative Horizontal is penalised by approximately 10% compared to the White British group. This outcome remains consistent across all three regressions (Tables 3.12, 3.15, and 3.16), with

Table 3.16

Graduate wage effects with strong negative/positive by horizontal mismatch and ethnicity – STEM graduates only

Ethnic Group	Horizontally Matched	Horizontally Mismatched
Blacks	-0.090** -(0.039)	
Chinese	0.093 -(0.090)	
Indian	-0.001 -(0.052)	
Other Asian Background	-0.055 -(0.069)	
Other Ethnic Group	0.016 -(0.045)	
Bangladeshi and Pakistani	-0.007 -(0.052)	
Other White	-0.011 -(0.042)	
Well – Matched White British	-	
Other White Horizontal		-0.043 -(0.061)
Strong Positive Horizontal		-0.041 -(0.085)
Strong Negative Horizontal		-0.102** -(0.052)
Control Variables		
Age	0.056*** -(0.005)	
Age Square	-0.060*** -(0.005)	

Female	-0.092*** -(0.012)
Tenure	-0.006 -(0.010)
Tenure Square	0.003*** -(0.001)
Horizontal Mismatch	-0.020* -(0.012)
Vertical Mismatch	-0.105*** (0.015)
Constant	**1.557*** (0.100)**
Observations	**5,237**
R-Squared	**0.3943**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places. The base category is the "Horizontally – Matched White British" group.

Tables 3.12 and 3.16 producing statistically significant results at the 5% level. This finding confirms the hypothesis that, generally, lower-performing ethnic minorities (excluding the Chinese and Indian groups), when coupled with horizontal mismatch, face significant penalties on hourly earnings regardless of their undergraduate field of study.

The majority of nominal ethnic minority results are negative, implying that, on average, ethnic minorities' hourly earnings are penalised compared to the White British group, consistent with previous regressions. While many of these results are not statistically significant, the trend persists. Considering the negative coefficients for ethnic minorities across previous regressions (with the Chinese group as an exception), it's highly likely that ethnic minorities face financial penalties compared to the White British group. This seems particularly pronounced when considering samples of non-STEM graduates.

Furthermore, the Black demographic once again experiences a significant penalty of 9%, resulting in the second-lowest hourly earnings compared to the White British group. This supports the finding that ethnic minorities generally receive inferior hourly earnings, as evidenced by consistently negative coefficients for the Blacks group across previous regressions, with statistical significance at the 5% level for Table 3.16 and the 1% level for Table 3.12. This strongly suggests that the Blacks group may have on average, the lowest degree classification/educational attainment, highest educational mismatch, highest discrimination severity, or a combination of these factors (Blackaby, Leslie, Murphy, O'Leary, 2002).

Among ethnic minorities, the Chinese group performs the best in terms of hourly earnings, consistently outperforming even the White British group. Conversely, the Blacks perform the worst, facing consistently significant penalties compared to the White British group and other ethnic minority groups. Additionally, Table 3.11 provides strong evidence that horizontally mismatched ethnic minority groups experience greater wage penalties compared to nominal ethnic minority groups. Overall, the Strong Positive Horizontal generally yields higher hourly earnings than the White British group, while the opposite holds true for the Strong Negative Horizontal.

3.6. Conclusion

3.6.1. Summary

The culmination of the extensive analysis offers valuable insights into the complex landscape of earnings differentials among ethnic groups in the UK labour market. Through a thorough examination of horizontal and vertical mismatches, subject choice disparities, and gender biases, a deeper understanding of the factors contributing to disparities in earnings outcomes has been achieved.

The analysis presented in Table 3.10 provides a nuanced perspective on vertical mismatch estimation using pooled Ordinary Least Squares (OLS) regression. By categorising individuals into under-educated, well-matched, and over-educated groups, a comparison of earning premiums associated with different education levels is possible. Notably, while ethnic minorities generally earned less than their White British counterparts, the Chinese group emerged as an exception, albeit not statistically significant. Surprisingly, the Other White group experienced a substantial earnings penalty, highlighting disparities even among Caucasian ethnicities.

Further exploration revealed that under-educated workers faced penalties across all ethnic minority categories, with the Bangladeshi and Pakistani group experiencing the largest penalty. Conversely, over-educated workers generally enjoyed a wage premium, consistent with prior research. However, some ethnic minority groups, such as the Black African and Other Asian Background groups, earned less than the well-matched White British group in the over-educated category.

The impact of educational mismatch on earnings was significant, with under-educated individuals facing the largest penalty. Our analysis underscored wage differentials across various ethnic groups, emphasising the presence of vertical mismatch between educational qualifications and occupational attainment.

In addition to vertical mismatch, the examination of horizontal mismatch revealed important insights into earnings differentials. The presence of horizontal mismatch, particularly among certain ethnic groups, significantly influenced earnings outcomes, highlighting the need for targeted interventions to address disparities within the labour market.

Moreover, the analysis of subject choice preferences among ethnic minorities elucidated a strong inclination towards STEM subjects, particularly among the Chinese and Indian groups. Gender biases in subject selection further contributed to disparities in earnings outcomes, with males showing a greater propensity for STEM subjects compared to females.

Overall, the comprehensive analysis underscores the multifaceted nature of earnings differentials among ethnic groups in the UK labour market. Moving forward, targeted policies aimed at addressing vertical and horizontal mismatches, promoting equitable opportunities, and reducing gender biases are essential for fostering a more inclusive and fair labour market landscape. Addressing these disparities can contribute to building a more equitable society where all individuals have the opportunity to thrive and succeed.

3.6.2. Policy implications

The policy implications of the findings in this study underscore a critical need for targeted approaches to mitigate both horizontal and vertical job mismatches among ethnic minority graduates, addressing wage disparities, and fostering an inclusive labour market. The data reveal significant earnings penalties associated with these mismatches, disproportionately affecting ethnic minorities and limiting the effective utilization of their skills. Addressing these gaps requires both systemic interventions and targeted strategies to dismantle the structural barriers affecting minority groups, particularly within the early stages of their careers.

Firstly, a comprehensive policy approach focused on reducing vertical mismatch is essential to enhance job alignment for ethnic minority graduates. Given that over- and under-education relative to job requirements result in wage penalties and job dissatisfaction, policies should aim to bridge these disparities by creating a more transparent pathway from education to employment. This could involve expanding internships, apprenticeships, and work-based learning opportunities, especially for minority students, to increase job-market readiness and enable smoother transitions into well-matched roles. Additionally, offering tailored career guidance and mentorship could assist ethnic minority graduates in securing positions that more closely align with their educational attainment, ultimately reducing vertical mismatch.

Horizontal mismatch also requires policy attention, particularly because certain ethnic groups experience more frequent and severe penalties when working outside their fields of study. Policies that promote skills portability and cross-sectoral flexibility would be beneficial, especially for those in highly specialized fields like STEM, where skill alignment with job roles is particularly challenging. This could include skill-recognition programs or reskilling initiatives that enable ethnic minority workers to leverage their expertise across diverse sectors. By broadening the applicability of skills acquired through higher education, ethnic minority graduates would have greater flexibility in navigating the job market, potentially mitigating the wage penalties associated with horizontal mismatch.

Educational institutions also play a role in addressing these mismatches, as they are integral to shaping career aspirations and field-specific expertise. A policy-driven collaboration between universities and industries could bridge the knowledge gap between academic training and labour market demands. Specifically, universities could enhance their career services by providing sector-specific counselling and organising events where students interact directly with employers from various industries. This would help ethnic minority students make informed decisions about subject choice and better prepare them for potential job market challenges, fostering greater alignment between education and employment outcomes.

Moreover, addressing gender disparities in subject choice is vital to closing the wage gap among ethnic minority graduates. Men are more likely to pursue higher-paying STEM degrees, while women are more frequently found in lower-paying sectors, contributing to earnings differentials. Initiatives to encourage diverse participation in STEM for both men and women within ethnic minority groups are crucial. Policies that support scholarships, mentorship programs, and STEM outreach for young women in ethnic minority communities could help create a more equitable distribution of talent across fields traditionally associated with wage premiums.

Employer accountability is another pivotal aspect of reducing job mismatch. Policies that incentivize inclusive hiring practices and fair wages for ethnic minority graduates are essential for reducing discrimination and creating a more supportive environment for skill alignment. The government could consider tax benefits or subsidies for employers who actively implement diversity programs, ensure equitable pay, and provide training that fosters cross-functional skills. This would not only help in reducing job mismatch but also encourage employers to value diverse skill sets within their workforce, promoting fairer employment practices overall.

Finally, structural changes to address systemic discrimination are fundamental to the success of these policies. Wage disparities linked to discrimination suggest the need for stronger anti-discrimination laws, with enforcement mechanisms that hold employers accountable for biased hiring practices or wage inequities. Additionally, increased data transparency on wages, hiring practices, and career progression could foster greater accountability and enable minority workers to advocate for fairer treatment in the labour market.

3.6.3. Limitations

While this study provides valuable insights into earnings differentials among ethnic groups in the UK labour market, it is essential to acknowledge several potential limitations that may impact the interpretation and generalisation of the findings. The analysis of the data is subject to limitations due to the sample size constraints, particularly in certain ethnic minority categories. For example, in the case for the Pakistani and Bangladeshi ethnic group, they were required to be grouped together throughout the entirety of the analysis. In the latter estimations, the Black ethnic groups were also required to be grouped due to a lack of sample size. If not grouped together, the standard errors would be large and would lead to a reduction in accuracy of the results due to inadequate representation of the population.

Moreover, there were an issue of respondents which did not fully complete the survey which are shown as "Missing Data" or "Does not Apply". Additionally, non-response bias may skew results as participation in the LFS is voluntary, potentially causing systematic differences between respondents and non-respondents. This diminishes the accuracy and completeness of the data has likely affected the reliability of the findings. For instance, many respondents did not complete the following surveys which resulted in the loss of valuable information and limited the potential depth of the analysis. Therefore, sampling bias occurs due to the reliance on a sample of households that might not fully represent the diversity of the population, leading to inaccuracies in labour market indicators if certain demographic groups are overrepresented or underrepresented.

While the LFS provides insights into key indicators, its scope may be limited in capturing aspects such as informal employment and job quality, posing challenges for comprehensive analysis. Additionally, a time lag between data collection and publication can limit the timeliness of information, affecting policymakers and researchers' ability to respond to rapidly changing labour market conditions. The aggregation of data at national or regional levels may obscure variations within specific demographic groups or geographic areas, necessitating additional data sources for fine-grained analysis.

3.6.4. Future research

Future studies, on the heterogeneity among ethnic groups regarding the impact of educational mismatches on earnings could provide valuable insights. This can deepen the understanding on the implications of vertical and horizontal mismatches in educational qualifications on earnings outcomes. Potential method could supplement quantitative analyses with qualitative research methods, such as interviews or focus groups. Investigating the policy implications of mismatch effects on earnings outcomes could also be useful by exploring the effectiveness of interventions aimed at reducing educational mismatches and improving earnings prospects for disadvantaged groups.

Additionally, examining the intersectionality of educational mismatches with other dimensions of social identity and conducting comparative studies across different national contexts could provide valuable insights into the labour market impact of educational mismatches and inform policy responses tailored to specific socio-economic contexts. Investigating employer perspectives on educational mismatches and their hiring practices could shed light on the demand-side factors driving mismatch outcomes in the labour market.

Finally, conducting longitudinal studies to monitor individual's earnings trajectories over time based on their educational attainment and mismatch status could offer valuable insights into how the effects of educational mismatches evolve throughout individuals' careers. Analysing data across multiple time points could help gauge whether these effects vary across different stages of the labour market, shedding light on the long-term consequences of mismatched education-occupation matches.

3.7. Appendix

Table A.3.1

Mean values of the control variables

Variables	Black African	Black Caribbean	Chinese	Indian	Pakistani and Bangladeshi	Other Asian Background	Other Ethnic Group	White British
Years of Education (Years)	20.64	18.52	21.45	20.61	19.85	20.14	20.47	18.21
Employed (%)	74.66	82.77	71.92	82.07	62.19	76.52	71.45	87.30
Real hourly earnings (£/hour)	12.44	13.26	16.23	15.81	12.00	12.78	13.67	14.18
Tenure (Years)	4.86	5.40	5.05	5.31	4.96	5.16	4.92	5.58
Highest Qualification Held								
Degree	47.97	30.75	70.65	57.40	45.66	45.47	48.80	32.32
Higher Education	12.55	13.02	4.69	7.69	6.08	8.07	7.71	9.99
A-Levels	11.58	21.99	6.31	10.30	14.59	12.40	9.52	24.27
GCSE	9.48	21.35	5.97	9.96	13.09	8.29	7.99	23.11
Other Qualification	13.63	8.62	8.28	10.08	12.25	19.08	18.81	6.19
No Qualification	4.78	4.27	4.10	4.57	8.34	6.68	7.18	4.11
Region of Workplace								
Tyne & Wear	1.05	0.17	2.99	0.86	1.42	1.43	1.19	2.76
Rest of Northern region	0.75	0.30	1.54	0.66	0.96	1.36	0.75	4.09
South Yorkshire	1.99	1.08	2.30	0.81	2.03	1.39	1.97	2.57
West Yorkshire	3.73	4.05	3.75	3.42	11.35	1.65	3.24	4.76
Rest of Yorkshire & Humberside	0.83	0.22	2.39	0.61	0.72	1.50	1.22	3.91
East Midlands	6.99	5.69	4.95	12.86	4.17	4.88	5.12	8.65
East Anglia	2.46	1.64	4.27	2.28	1.91	3.74	3.06	5.23

Central London	12.00	14.75	13.82	11.20	8.57	10.72	15.41	3.76
Inner London (not central)	16.00	16.90	10.15	6.75	10.05	9.21	11.39	2.02
Outer London	15.00	17.59	8.53	17.57	10.91	20.04	13.88	3.26
Rest of South East	17.00	12.63	17.49	17.46	13.84	24.00	19.34	20.48
South West	3.00	3.00	6.48	3.70	1.91	6.31	6.05	11.07
West Midlands Metropolitan	5.00	13.93	5.00	10.00	12.00	3.00	4.00	3.90
Rest of West Midlands	2.00	2.03	3.50	3.15	2.75	2.39	3.24	5.15
Greater Manchester	7.00	3.45	6.40	3.70	11.00	2.64	4.15	5.11
Merseyside	1.00	0.34	2.56	0.64	0.67	0.84	1.00	2.33
Rest of North West	1.00	1.03	1.45	2.72	4.57	2.46	2.40	5.12
Wales	1.00	0.73	2.05	1.15	1.42	2.13	2.22	5.58
Strathclyde	0.00	0.00	0.00	0.03	0.00	0.11	0.00	0.03
Rest of Scotland	0.00	0.09	0.09	0.09	0.17	0.11	0.19	0.13
Outside UK	0.00	0.04	0.09	0.06	0.03	0.11	0.06	0.08
Managerial Status								
Manager	16.31	21.47	25.60	26.68	18.44	19.60	20.19	28.44
Foreman or Supervisor	15.29	12.51	10.49	11.29	11.00	14.61	11.39	11.64
Not Manager or Supervisor	68.40	66.02	64.00	62.00	70.58	66.00	68.00	59.92
Marital Status								
Married	57.70	44.03	71.08	78.38	71.92	76.26	72.00	71.32
Non-Married	42.30	55.97	28.92	21.62	28.08	23.74	28.00	28.68
Number of employees at workplace								
1-10	14.51	13.97	22.95	13.38	23.02	19.49	19.03	18.00
11-19	7.99	7.59	5.03	6.28	7.73	6.64	8.00	8.00
20-24	4.04	3.84	4.95	3.51	3.42	4.84	4.00	4.26
Don’t know but under 25	2.96	2.29	1.37	1.50	2.75	2.24	2.12	1.45

25-49	11.53	11.86	8.36	10.65	10.68	12.04	11.05	13.72
50-249	21.81	23.00	17.75	22.22	21.71	19.34	21.31	24.61
250-499	6.11	7.93	6.14	8.63	6.00	6.31	6.74	7.53
Don't know but between 50 and 499	5.11	3.00	3.07	3.53	3.85	3.00	4.06	2.76
500 or more	25.93	26.39	30.38	30.30	20.00	26.20	23.81	19.93
Number of children								
0	37.66	53.94	56.57	42.58	29.80	41.14	45.79	58.16
1	20.97	24.37	20.14	25.74	19.74	23.49	22.73	18.61
2	23.38	15.79	20.00	24.73	26.13	27.95	22.38	18.26
3 or more	17.99	5.90	3.50	6.95	24.33	7.42	9.09	4.98
Observations	**3,617**	**2,319**	**1,172**	**6,545**	**3,454**	**2,725**	**3,205**	**217,102**

Chapter 4

An Investigation of the Returns to Generic Skills in the UK Labour Market

4.1. Introduction

Since the 2008 financial crisis, the UK economy has experienced a relatively poor productivity performance – this can be partially explained by the "Productivity Puzzle" which has caused the real wages to remain below the pre-crisis peak. The poor productivity is a long-established phenomenon, with the leading issue identified by an underinvestment in skills which has restricted economic growth (Industrial Strategy, 2017). Therefore, the UK government and policymakers have extensively focused on improving the generic skills of the UK labour market, especially to accommodate the significant changes in the modern workplace (Darrah, 2013; Appelbaum et al., 2000). Specifically, the UK government has concentrated on certain generic skills such as communication skills, numeracy, information technology skills, problem-solving skills, and skills to enhance teamwork ability. Additionally, in more recent times, the UK as well as globally have increased their investments in the skills attributed towards computing in seeking to adapt to more technological advancements (Borghans and Weel, 2004). The idealised aim is for the UK labour force to acquire skills to increase prosperity and in turn improve the living standard across the whole country.

In the previous chapter, it is found that the primary determinant of an individual's wage is intricately linked to their level of education, a widely employed metric in economic research for comprehending prevailing trends. However, relying solely on education as a metric may present an incomplete picture, as it fails to consider the level of skills possessed by the worker. It is noteworthy that education primarily gauges the characteristics of the worker rather than the specific tasks associated with the job (Handel, 2020). Interestingly, reports such as by Velsco (2014) suggest that a substantial number of workers perceive a misalignment between the required education level for their occupation and their actual educational attainment, indicating the existence of vertical and/or horizontal educational mismatches. Quintini (2014) suggests that this phenomenon may arise because education is often perceived as a certificate, signal, or screening mechanism to assess occupation accessibility, considering individual characteristics such as motivation or compatibility in the labour matching queue. Consequently, it becomes imperative to not only assess the education level but also the prevailing skills in the UK labour market.

In the modern workplace, many industrialised countries have acknowledged the significance of improving certain skills that were perceived as deficient in certain segments of the workforce. Governmental bodies and interventions such as the International Labour Office (2010) in Geneva have enacted policies emphasising a range of essential skills. These include communication capabilities, numerical proficiency, expertise in information technology, problem-solving aptitude, and the skills necessary for collaborative work to accommodate a skilled workforce for strong, sustainable and balanced growth.

The Skills and Employment Survey (SES) dataset will be utilised to evaluate how the generic skills impact earnings. The conceptuality of skills is increasingly prominent considering the education levels have been rising overall over a sustained period due to reasons such as changing job demands from continuous rapid technological advancement. According to Handel (2000), the increase in skills is not just for higher paid jobs, rather it is throughout the UK labour market, even in occupations whereby the skill development is deemed to be relatively stagnant or in less skilled occupations such as taxi drivers. In general, education levels have increased relatively uniformly throughout the quantile of the UK labour force, thus using education independently is not a proficient proxy to grasp the required education or the skills required explicitly for an occupation – this analyses the worker side as opposed to the job side measure.

In this chapter, Principal Component Analysis (PCA) is employed so reduce 35 sub-skills into six generic skill indices to enhance the interpretability of variables. Following the derivation of these generic skills, econometrically, the methodology builds on Dickerson and Green (2004) by integrating a Quantile Regression (QR) approach. Thereafter, estimations are conducted using an Ordinary Least Squares (OLS) method and a hedonic wage equation to derive the coefficients representing the shadow prices associated with attributes or skills.

Recognising this limitation, Koenker and Bassett (1978) introduced a solution known as the Conditional Quantile Regression (CQR) to mitigate the drawbacks of OLS given the prevalence of outliers and long-tailed distributions in wage equations. Moreover, OLS often omits important information such as understanding the effects across the wage distribution since it focuses on the conditional mean. Therefore, building on the methods of previous studies, the wage equation is estimated with various quantile points, as outlined in Chapter 3. This will allow for a better understanding of the relationship between the generic skills outside of the mean of the data which is useful to discover outcomes that have a nonlinear relationship with the explanatory variables.

The CQR approach estimates the conditional quantile of the distribution, allowing for the investigation of statistical variations in parameters across different quantiles of the wage distribution. The advantage of the CQR includes robustness against heteroskedasticity and outliers, greater efficiency than OLS when confronted with non-Gaussian error terms and detailed insights into distribution differentials by computing multiple regression curves corresponding to the specified earning quantiles within the distribution.

Despite the benefits of the CQR, there are circumstances which cause limitations. For instance, changes in the generic skills at various wage distributions and the distribution conditional on all observed characteristics can affect and instigate misinterpretation of the coefficient of these variables. As such, the distribution of the wage premium at a specific θ^{th} quantile is contingent on the wage distribution of the specified generic skill variable and all observed characteristics. Therefore, the inequality generated by the generic skill can influence the distribution of inequality of the wage premium which may hinder the result of the estimation to be either more or less than the actual value.

To resolve the limitations of the CQR, Firpo, Fortin and Lemieux (2009) implemented a regression approach known as the Unconditional Quantile Regression (UQR). The distinction between these two approaches lies in the modelling strategy. In the CQR (Conditional Quantile Regression) method, the estimation of the wage premium considers the influence on generic skills and observed characteristics within the specific group of interest. Conversely, the UQR (Unconditional Quantile Regression) method estimates the impact on the wage premium at a particular θ^{th} quantile across all observations in a group, without conditioning on generic skills and observed characteristics (Fulden, 2021). Therefore, the UQR can analyse the relationships between the generic skills and produce a more extensive result. The methodology will further explain how the outcomes of a UQR are more interpreTable and generalisable.

Finally, further contribution will be made by investigating whether the Non-White ethnic group exhibit greater or lower job generic skill requirements than their counterparts – if so, to what extent and how does the generic skills disparity affect the pay between the two groups. According to Brynin and Guveli (2012), there is a significant pay gap between the white British workers and the ethnic minority groups which suggest that discrimination might be a factor. Discrimination appears to manifest in two forms: initial barriers preventing non-white individuals from accessing well-paid occupations, and within-job discrimination leading to lower pay for non-whites compared to their white counterparts in similar occupations.

4.2. Literature review

4.2.1. Importance of skills

Individuals pursue training to enhance their skills, aiming to improve their earning potential. According to Attewell (1990), skill is intrinsically difficult to define and measure. The traditional method to measure the relationship between skills and wages have been conducted by using the returns of education. For instance, Becker (1964), introduced the concept of human capital and the internal rate of return of human capital. He also demonstrated how internal rate of return of human capital can be employed to choose among alternative investments in human capital. Following, Mincer (1974), implemented the Mincer earnings function which investigates the human capital earnings by estimating the effects of schooling. However, employers stress that a significant number tend to concentrate excessively on specialized skills, neglecting the crucial value of generic skills such as literacy, communication, teamwork, and problem-solving (Becker, 1975). Despite their importance, these generic skills are often overlooked.

Previously, skills are understood to include four fundamental components: underpinning knowledge, specific skills, ability and other characteristics such as motivation and dependability (Form, 1987). From Mincer's (1989) findings, there is strong evidence that the four components correlate with educational attainments, although it is difficult to conceptualise which parts of skills are rewarded, and the outcome is likely to differ from various occupations. Kenny et al (1979) have added that in terms of econometrics, the estimates of returns on education utilising Mincer's (1975) method may exhibit an inherent upward bias. This bias arises due to the unobservable nature of innate skills, with educational attainment and skills expected to exhibit a positive correlation. Consequently, this introduces greater uncertainty when attempting to gauge the impact of skills on wages through educational attainment as a proxy. Arrow (1973) and Spence (1978) argue that systemic shifts in the production process have resulted in changes in the demand for specific types of labour. Therefore, relying on education as a proxy can be problematic, as it implies that employers reward qualifications and schooling, potentially overlooking the actual skills individuals possesses and exercise in their workplace.

Recent studies on skills have shifted focus from the "supply" perspective to investigating the "demand for skills." According to Green (1998), the absence of standardised assessment of specific generic skills during workplace makes it challenging for employers to appropriately recognise and reward skilled workers. Notably, studies that measures skills utilisation in workplace are predominantly focused on UK data, as evidenced by Ananiadou et al. (2004), Ashton et al. (1999), Dickerson & Green (2002), Felstead et al. (2002), Green & McIntosh (2007), Green et al. (2002), Spenner (1990), and Walker & Zhu (2007). For instance, Dickerson and Green (2002) utilised the 1997 and 2001 UK Skills Survey to examine changes in various skills, including computing and generic skills like literacy, numeracy, high-level communication, planning, client communication, horizontal communication, problem-solving, and checking skills. Their findings revealed positive wage premiums for computing and high-level communication skills. Ananiadou et al. (2004) examined several UK skills training studies, encompassing wage and employability as dependent variables. They concluded that enhanced numeracy and literacy skills significantly contribute to individuals' earnings and employability. This effect remains robust even after accounting for demographic factors, personal attitudes, and soft skills.

Additional research on skills as exemplified by Rotundo and Sackett (2004), examined the comparative influence of generic and specific skills on wages. Their findings indicated that variations in wages were primarily attributable to generic skills rather than job-specific ones. They explained that “investing in the development of a diverse range of cognitive skills/abilities yields the greatest returns” (p. 145). In a study between the periods of 1993 and 1996–1997 using enterprise surveys in Ireland, Barrett and O'Connell (2001) investigated the impact of skills. They demonstrated that generic training had a positive impact on productivity (measured by wages), while the influence of specific training on wages was not significant. The positive effect of generic training on wages persisted despite controlling for occupations, corporate restructuring, firm size, and the level of human capital.

However, Barrett and O'Connell (2001) research contradict the human capital theory's proposed by Becker (1975). since employers only see the benefits of specific training as it will augment productivity of the current job. According to human capital theory, employers are expected to perceive the advantages of specific training solely in terms of its potential to enhance the productivity of the current job. Consequently, the theory suggests that employers would not invest in general training, as it might lead to an increase in productivity for other employers if workers are subsequently poached. However, empirical findings deviate from these expectations as employers endorse general training and recognise the value of generic skills, thus rewarding them accordingly (Barrett and O'Connell, 2001). In response to this disparity between theory and reality, economists like Riveros, Bouton, and Mundial (1991) argue that the efficiency wage theory and psychology of contracts are commonly employed to interpret this unexpected outcome.

Florida et al. (2012) investigated the relationship between skills and wages in the US. They utilised data from the US Occupational Employment Statistics of 1999 and 2008, along with the O*Net database. The primary focus of Florida et al.'s study is to determine the increasing significance of cognitive and non-cognitive skills in the US economy. Additionally, they examine whether these skills exhibit a spatial distribution, with cities, and their sizes, being more likely to benefit from cognitive and non-cognitive skills, while rural areas may rely more on traditional industries where physical skills are more relevant. The findings of Florida et al. reveal a positive correlation between analytical and social intelligence skills and higher wages, particularly in larger cities where these skills are more prevalent. In contrast, studies indicate that physical skills have a negative impact on wages overall, and there has been a declining demand for physical skills in industrialised economies (Autor & Handel, 2013; Autor et al., 2010; Green, 1998; Felstead et al., 2002).

Research investigating the relationship between generic skills and wages is scarce in Asian countries. However, Ramos et al. (2013) showed in Singapore there are significantly positive and strong relationships between wages and generic skills such as leadership, planning, literacy, and problem-solving skills. Their study utilized the 2011 Skills Utilization Survey in Singapore project, conducted under the Institute for Adult Learning (Sung et al., 2013). In line with findings in the UK and US, Ramos et al. observed that physical skills do not significantly contribute to wage effects. This might be explained by Singapore's geographical location, surrounded by less developed countries with relatively low labour costs. Similar to Dickerson and Green (2004), the research found a negative relationship between numeracy skills and wages, although statistical significance was not achieved in either study. Ramos et al. concludes that the lack of significance for numeracy and other skills does not necessarily imply an absence of a significant relationship with job incomes. Instead, skills such as numeracy may be expected, like our ability to stand and walk. Attewell (1990) asserts that although most individuals can perform basic calculations, they are no longer rewarded for these skills, even though they may be crucial for certain job tasks.

4.2.2. Theoretical background

The foundation of skill acquisition and wage determination in labour economics is heavily influenced by Becker's (1964) Human Capital Theory, which distinguishes between general and specific skills. General skills are transferable across firms and industries, providing employees with broader employment opportunities. However, since these skills can be utilized by multiple employers, firms are less incentivized to invest in general skill training. Specific skills, on the other hand, are tailored to a particular firm or industry, benefiting the current employer directly, making firms more willing to invest in their development. This distinction has long been a cornerstone of understanding labour market outcomes, wage growth, and training investments, especially in cases where firms assess the return on their human capital investments.

However, subsequent research, notably by Stevens (1994), expanded on Becker's framework by introducing the role of labour market conditions in shaping firms' training decisions. Stevens argued that factors such as job search costs, unemployment rates, and the competitive environment in the labour market influence firms’ strategies on skill development. In tighter labour markets, where competition for workers is fierce, employers may invest more in skills that are specific to their needs in order to retain workers. On the other hand, in periods of higher unemployment, firms may underinvest in skills, expecting that they can recruit externally at a lower cost. This perspective adds nuance to the discussion of how external market conditions shape internal human capital decisions.

Further theoretical advancements came from Acemoglu and Pischke (1999), who challenged the clear-cut distinction between general and specific skills. Their work highlighted that, under imperfect labour market conditions, firms may also invest in general skill training. In real-world labour markets, where there are information asymmetries or mobility constraints, employers may still benefit from investing in skills that are technically transferable, as employees are less likely to leave for alternative jobs. This blurring of the boundary between general and specific skills suggests that firms’ training investments are more complex and context-dependent than originally envisioned by Becker. In markets with low labour mobility, firms can capture some of the returns on general training, thereby treating it similarly to specific skill investment.

These theoretical models provide a critical lens through which to view the UK labour market. Skill mismatches, as highlighted in earlier sections of this thesis, persist across many sectors and regions, leading to variations in wage returns. The impact of technological advancements, particularly in IT and computing, has further complicated the traditional general-specific skill dichotomy, as these high-demand skills often exhibit both general and specific characteristics depending on industry needs. Additionally, the UK's regional labour market disparities—particularly with respect to high-skill worker retention—align with Acemoglu and Pischke’s predictions that firms in certain regions may still invest in ostensibly general skills due to mobility frictions. Understanding these theoretical advancements helps contextualize the empirical findings and sheds light on the nuanced dynamics of skill development and wage determination in the UK.

4.2.3. UK generic skills

The measurement of generic skills of the UK workforce is invaluable to understanding the levels of skills required for the wide range of jobs across all industries in the UK. Early papers such as those by Dickerson and Morris (2019) suggest that work-related skills increase earnings in the UK and argue it may even be more important than educational qualifications, especially within occupations which demand high analytical and interpersonal skills. This proposes that UK policymakers should promote both analytical and interpersonal skills while relaxing the focus on physical skills. Estimation using the decomposition method shows that between 2002-2016, there is a positive and significantly increasing returns over time with both analytical skills and interpersonal skills. However, physical skills are found to be significantly negative and there is a declining usage of physical skills in the UK.

Dickerson and Morris (2019) implemented a Mincerian earnings regression and found that the returns to analytical skills and interpersonal skills are not only positive but also increases over time which implies that the demand for these skills is expanding greater than the growth of their application (vice versa for physical skills). This suggests that the UK labour force demand for cognitive skills is likely still increasing, although this does not signify whether the increase in skills is spread uniformly across the UK.

On the other hand, economists such as Greenan, Kalugina and Walkowiak (2014) argue that the labour force of the EU-15 from the EWCS (European Working Conditions Surveys) throughout 1995-2005 has had an increase in the physical strain as well as work intensity, whereas the work complexity has decreased. The overall result remained the same despite controlling the variables including worker age, gender, type of employment contract, supervisory status, computer use etc. This finding is contradicting the papers previously mentioned such as by Dickerson and Green (2004) since they found that all the generic skills for the UK labour force increased except for physical skills. Indeed, these two papers investigate different areas whereby the former examines the working conditions, and the latter observes the changes in generic skills.

However, one would expect that if physical skill were overall demanded less, the physical strain would decline, particularly as there is a constant increase and improvement in technological advancements. Greenan et al (2014) acknowledge the conflict of expectation and have stated that institutional explanations face the difficulty of understanding how Germany, Italy, Spain, and the UK experienced the greatest decline in work complexity. They concluded, "Our statistical analysis identifies a complexity paradox and leaves it unresolved".

More recent, and updated papers such as those by Dickerson and Morris (2019) had similar results to the earlier paper by Dickerson and Green (2004). They examined the development of occupational composition and skills in the UK over the period between 2002 and 2016 by focusing on three explanatory variables referred to as the three indices of skills: analytical, interpersonal, and physical skills and investigated how these controls affect the conditional returns to earnings of workers. Consistent with previous research (e.g., Dickerson and Green, 2004), the UK labour force's analytical and interpersonal skills increased over time whereas physical skills declined. During the study period, the increase in returns for interpersonal skills for females was lower than their male counterparts, this is particularly evident post-2010. This finding is consistent with the research for the US by Deming (2017) and the research for Sweden by Edin et al (2022). As a result, there is no indication that a greater cognitive skill by the UK labour force reduces returns; however, the returns to physical skills are significantly lower than zero during the period of study.

4.2.3.1. Relationship between graduates and generic skills

There are several main instigators for the heterogenous development of generic skills for students that are studying in higher education including: during educational context, individuals' willingness, culture and background, degree/disciplinary course, and industry training. Firstly, even though the educational framework can enhance one's generic skills, there is usually a consensus among university academics and researchers that universities and higher educational frameworks should focus prior on the teachings of the specific field of interests as opposed to the development of students' generic skills (Bennett, Dunne, and

Carré 1999; Star and Hammer 2008). In most cases, universities advocate academics to research and provide publications since this increases the university's rank and credibility, thus emphasis is instigated upon research activities rather than teaching (Drummond, Nixon, and Wiltshire 1998; Jung and Chan 2017).

Second, is the student's exercise to accommodate extra curriculum under their expense in time and money. According to Dunne, Bennett, and Carré 1997; and Arevalo et al. (2010), students under their influences, tend to neglect the importance of the development of generic skills. Students who are approaching graduation tend to concentrate on their academic endeavours, while younger students are less inclined to value the importance of generic skills.

The third factor which determines the development of generic skills is the various cultures and backgrounds of an individual. For instance, the focus of examination is even more profound in Asian cultures which usually fully prioritise academic attainments at the opportunity cost of complete personal development as developing generic skills is often deemed as 'time-consuming' and, thus overall, not an optimal trade-off (Leung, Leung and Zuo, 2014). This is supported by Chan and Fong (2018), every discipline has its characteristics and culture which incorporates a spectrum of specific academic knowledge and generic skills.

Fourthly, academics suggest that depending on the degree chosen, some courses tend to focus primarily on disciplinary knowledge as opposed to the development of generic skills (Leckey and McGuigan 1997). For example, degrees that tend to be specialised such as accounting and engineering almost always primarily focus on the academic content. This effect of focusing on disciplinary knowledge is even more profound under circumstances whereby the objective is to obtain accreditation, even though possessing several generic skills is partially the requirement to achieve the accreditation. As a result, the majority of universities minimise the inclusion of generic skills while teaching and students themselves tend to prefer to concentrate on technical content to strive toward their professional accreditation (Chawner and Oliver, 2013).

Moreover, Jones (2009a and 2009b) examined the generic skills development within a cross-disciplinary setting which comprised five-degree courses including economics, history, medicine, law, and physics. She concluded that generic skills were taught contrastingly as certain degree courses more focused on particular generic skills than other skills, and vice versa. Therefore, depending on which degree course an individual chooses, there is a significant impact on how and which generic skills are developed. As a result of this finding, Potrac and Jones (2009a and 2009b) proposes whether generic skills should be implemented in higher education and the focus of the specificity of generic skills development is pre-determined with the dependence on the discipline chosen by the student.

Therefore, while various degree courses focus on different generic skills, the method of teaching also tends to be distinctive depending on how the technical material and generic skills development are presented in the most suitable approach. According to Ballantine and Larres (2004), engineering degree tends to undertake project-based workshop so that students can experience working scenarios that includes challenges such as technical problems. These scenarios are often performed in teams whereby challenges require knowledge in the field as well as various generic skills such as creativity, problem-solving and teamwork. In business degrees, the curriculum often includes theoretical as well as practical teachings. For example, the inclusion of case studies enables business students to understand the knowledge and generic skills to solve real-life problems when they encounter future occupations (Boyce et al, 2001). These are arbitrary examples for engineering and business degrees and are bounded by the facilities and capabilities of the classrooms and workshop settings, thus the generic skills obtained from a particular degree are difficult to transfer to a different degree.

Additionally, there are methods for individuals to develop generic skills through training within an industry. However, there is an extent of uncertainty that the returns to industry-specific training are not successful since the worker may fail to find employment in the industry that he/she is trained for. This is a loss in investment for both the employees and employers since the workers suffer a significant wage loss when they switch industries and the business will endure a loss from the investment training cost (Carrington, 1993; Neal, 1995; Weinberg, 2001). As a whole, businesses will have a larger drawback if this scenario happens as workers may benefit from developing not just skills that are specific to the industry, but also the various generic skills that can be applied in future occupations. Conversely, businesses have minimal incentive to provide potential workers with supplementary training than required to be productive in the firm.

4.2.4. International measurements and generic skills

In contrast to the UK datasets, the US primarily utilises the US Dictionary of Occupational Titles (DOT) or the newer sequel O*NET (Occupational Information Network) to measure tasks and generic skills (e.g., Abraham and Spletzer, 2009) which captures a greater scope of individual generic skills than the dataset in the UK. Alike the paper by Green and Dickerson (2004), the disadvantage of using the SES dataset is that the measure of generic skills is derived from job requirements as opposed to the actual skills possessed by the employee. Using occupational classification and required skills for the occupation is a poor proxy to measure skills as the knowledge or abilities learnt during education may be forgotten or not applicable in the modern era (Dickerson and Morris, 2019). It is proclaimed that while “skills” have been a foremost matter for UK policymakers (e.g., UKCES, 2009, 2010, 2014), there are scarce imperfect measures of skills which can be used to examine employment (Musset and Field, 2013). To facilitate the data collection for individuals' generic skills and computer competence in the UK, it would be exponentially more challenging and expensive to gather the information.

Autor, David and Dorn (2013) investigated the dynamics of the changing skill demands and wage structure of low-skilled, educated and low-wage employment in the US labour market. The overall results of the generic skills from this research are consistent with the findings by Dickerson and Green (2004). As such, like the UK, the US workforce's generic skills have generally developed, while the physical skills have diminished in later years which is as expected.

In terms of technological advancements, Autor, David and Dorn (2013) hypothesise that computerisation should substitute low-skilled occupations which incorporate routine tasks, while highly educated workers would be in occupations that require creative, problem-solving and coordination skills. Their result illustrates that due to the continuous price depreciation of the general computer technology, the wage paid for routine tasks has subsequently been spiralling downwards. This has diverted low-skill workers to service occupations which are complex to automate since some procedures rely extensively on flexible interpersonal communication and direct physical management. Autor, David and Dorn (2013) suggest that if the good and service output does not possess a comparable production substitute, then the increase of information technology for routine tasks would induce an increase in wages and employment in low-skilled occupations. Therefore, wages are not only affected by the level of skills a worker possesses but also the ability to substitute a routine task is an important factor as Baumol (1967) originally identified; a model of unbalanced technological progress concerning skill demands and wage structure.

Australia's approach (Badcock, Pattison, and Harris, 2010) to measure the level of generic skills is similar to the SES questionnaire produced by Alan Felstead and colleagues whereby there is a scale of five measurements to indicate the level of a specific generic skill. Badcock, Pattison, and Harris (2010) utilise the graduate skills assessment (GSA) to examine an individual's level of generic skills which includes critical thinking, interpersonal understanding, problem-solving and written communication. In this study, they focused on three disciplines: arts, science and engineering in an Australian university and found that there is a distinct difference in generic skills scores in various disciplines. This finding is consistent with Potrac and Jones (2009a and 2009b) as it suggests that individuals are more developed in certain generic skills than others depending on the degree course chosen. This is as expected since the generic skills required in an art occupation would be substantially different from an occupation in science. They observed that engineering students' critical thinking, interpersonal understanding and written communication scores were significantly lower than those of the other two disciplinary groups. This is emphasised by Gimenez (2012) as he found that students within the disciplines of nursing and midwifery which are considered similar have a significant effect on academic writing during their undergraduate course.

In Chan and Fong (2018) research, they explored the implications of disciplinary differences between first-year engineering and business students in the Southeast Asian region of Hong Kong. They examined the students at a Hong Kong University to observe whether there are differences in the incorporation of generic skills development within their curriculum context.

Consistent with Badcock, Pattison, and Harris (2010), business students scored higher on the importance of self-management, interpersonal and communication skills compared to their engineering counterparts is evidence that students who undertake certain degree courses are more attentive to the importance of developing generic skills. However, engineering students scored higher on the importance of problem-solving and critical thinking skills which is attributed to the ABET, HKIE and CEF accreditation. This suggests that students in different degree courses are more prone to research the specific generic skills which are required for their accreditation to suitably equip their identity to enhance their entrance into the industry (Jackson, 2016).

However, some studies have found that a range of generic skills can only be grasped and developed within disciplinary contexts (Pattison, and Harris 2010; Sweetman, Hovdhaugen, and Karlsen 2014). Universally, higher education students must be taught generic skills within their disciplinary context since it is challenging for students to understand the benefits of developing these essential skills which will be applied to future occupations. Therefore, arguably international universities and higher educational entities should incorporate generic skills integral via their course curriculum.

4.2.5. IT and computer skills

The advancement of computing skills and ICT has rapidly become a great economic driving force since the late twentieth century and the twenty-first. The UK experienced a vast expansion in the domain of Information and Computer Technology (ICT) during the end millennium and has become an increasingly important topic (Dolton and Makepeace, 2004). There has been a substantial increase in ICT investment and the relative concentration of the total investment as a sector increased exponentially. According to Dolton and Makepeace, ICT became a significant instigator of the expansion of the UK economy and has generated an upsurge in labour productivity.

"The level of computer use" is a valuable indicator to supplement the examination of the varying complexity of an individual's computing skills. Estimations derived by Dickerson and Green (2004) have deduced four main findings. Consistent with previous studies such as

Felstead, Gallie, and Green (2004), there is an overall increase in most generic skills (except for physical skills). In particular, computing skills are found to develop rapidly. Also, there is a positive wage premium with high-level communication skills and computing skills. More specifically, the more advanced and complex implementation of computers attribute to a greater premium as opposed to a more basic practice.

Moreover, Dickerson and Green (2002) found that keeping other variables constant, a worker which requires the use of computers in her/his job experiences higher pay than their counterparts who do not require to use of computers in their workplace. Finally, the generic skills in jobs possessed by men and women are considerably diverse as expected since many men (women) work in different occupations than women (men), thus requiring contrasting skills. Despite the discrepancies in the skills between the genders, Dickerson and Green found that there is no unexplained gender pay gap which is usually due to discrimination. However, Dickerson and Morris' (2019) paper found that the returns to interpersonal skills for females were statistically lower than for males. (IT and computing skills section)

Between the 1980 and 1999, there was significant capital investment for computers and software which attribute a growth rate of 30% and 32% respectively (Outlon, 2001). This finding can be reinforced by Colecchia and Schreyer (2002) as they calculated that ICT equipment and software of non-residential investment for the UK doubled between 1980 and 1990, while it tripled between 1980 and 2000. The growth of ICT was so profound that Outlon claimed that irrespective of the minimal share in GDP, the domain of ICT accounted for 13% of the UK economy's output growth between 1979 and 1989 whereas 21% between 1989 and 1999. Hence, the focus is on investigating the growth of computing and ICT, along with an exploration into the development of individuals' computing skills – does this significantly affect the wage structure?

Other papers such as Dolton and Makepeace (2004) looked at the returns of computer use amongst three different groups between the period of 1991 and 2000. The three groups include individuals who used computers in both periods (stay), only the first period (leave) and only the second period (enter). The result showed that the rate of return of computing over time does not increase uniformly, conversely, the magnitude of return from computing differed across individuals. The estimates showed that male leavers achieve an earnings premium of 9%. Since Dolton and Makepeace excluded fixed effects and incorporated a diverse set of explanatory variables, it is not possible to attribute this to unobserved ability. Interestingly, the estimate for men exemplifies that the impact of computer use remained constant in both periods for stayers and enterers, although the outcome was zero for leavers in 2000. In contrast, women received an earnings premium of between 10% to 12% with computer use. Thus, there is a significant earnings premium associated with computer use, though the premium is slightly higher for women than men.

Previous papers have found that earnings are generally higher for workers who require computer usage in comparison to occupations that do not require computers. In Krueger's (1993) research, he found that workers who use computers in their jobs earned 10% to 15% more than their counterparts. It was advocated those technological changes have increased wage for computer-using workers since the spread of computer at work has generated a demand for this specific skill. Considering highly educated workers are more likely to use computers at work, between one-third and one-half of the increase in the rate of return to education is due to the expansion of computers.

However, there are some studies such as DiNardo and Pischke (1997); Entorf and Kramarz, (1997); Haisken-DeNew and Schmidt (1999) which challenge Kruger's theory. There are fixed effects estimations that indicate that the wage increase from a computer usage occupation is considerably diminished when the correlation of computers with unobserved individual heterogeneity is included in the model (Dickerson and Green, 2004). Entorf and Kramarz (1997) explicate that only a minimal wage gain is achieved from operating computers in jobs. Estimates in the 1980s, the increase in wages from computer usage in jobs did not exceed 5% after 6 years and was even lower in the 1990s (Entorf and Kramarz, 1998).

The main concern concerning Krueger's finding is that since the workers who use computers in their jobs are in general higher in both education level and skills, their productivity would be greater regardless of whether computers are used in the workspace. As a result, these workers would naturally receive a wage premium due to their greater ability in comparison to their counterparts. Moreover, the dataset used by Krueger has very restricted information about the use of computers and computer skills. For example, the explanatory variable used to examine the intensity of computer use in jobs is derived via "whether or not a computer is used at work". Consequently, there is no accountability for the level of intensity of computer use or how essential the usage of computers is associated with specific jobs.

Other economists such as Bisello et al (2019) revised the findings on the EU-15 by using the more recent 2015 EWCS dataset and the results are mainly consistent with Greenan et al (2013). However, there is evidence of a contradictory effect concerning the computerisation hypothesis in terms of social and routine task content. Despite the large-scale increase in computerisation in the past two decades, there is an even greater increase in the repetitiveness and standardisation of the work process as well as a decline in social task content. Thus, computers seem to be simultaneously displacing labour into social tasks and decreasing the social tasks in the remaining jobs. Other measures such as cognitive complexity and problem-solving indicate no change over the two decades, while both the demand for interpersonal skills and physical skills declined. Surprisingly, task repetitiveness and monotony increased whereas the most substantial changes include a decrease in machinery use and an increase in computer use as expected.

Consistent with previous papers such as Dolton and Makepeace (2004), they found that the expansion in computer use has had the greatest change over the past two decades by a considerable margin. This finding contradicts the status quo hypothesis as this would imply that the advancement of IT has insignificant effects on jobs and earnings. Examination will be conducted to determine whether the regression analyses follow the trend where significant growth in IT use translates into relatively modest or even counterproductive job characteristics.

The supply of a specific skill/skill is significantly determined by the demand for the skill/skills. Bartel et al. (2003) examines whether the implementation of new IT increases the demand and supply for certain skills. For example, the exercise of sophisticated technology could increase the level of skills required which usually has a direct relationship with the demand for computer skills. Moreover, the skill demand could be increased indirectly since the issuing of IT tends to require two types of operators – those who instigate the programming of the machine and those who run the programs of the machine which generally possess less knowledge of the machining functions. If businesses adopt new machines which are complex to operate, the overall skill demand should increase.

Bartel et al. (2004) later assessed the differences in demand for various skills between the US and the UK plants when subjected to an increase in IT use i.e., Computerised Numerical Control (CNC). In recent times, there has been a significant increase in the demand for computer skills, programming skills, problem-solving skills and engineering skills due to a direct effect of an increase in IT usage. They found that the UK had a greater increase in demand for computer skills and programming skills compared to the US when there was an increase in IT implementation via CNC machines. Therefore, this particular case study suggests that the UK is more susceptible when there is technological increase, development or advancements. It can be argued that in this field, the UK is more inclined to ensure the labour supply is well equipped in their respective area. The demand for skills in the US economy is continuously growing due to technological advancements which have in turn increased wage inequality (Bartel et al., 2004).

Overall, the invention of new technology requires organisations to incorporate higher levels of cognitive skills, flexibility skills and autonomy than in conventional labour settings where previously the production process was relatively constant with limited variables. According to Bartel and Lichtenberg (1987), cognitive skills must be proficient to adapt to changes, such as when implementing new technologies. In recent years, the labour force IT skills have substantially improved due to the labour demand by the majority of modern firms – thus, there is now a substantially greater number of highly skilled IT labourers due to the demand by firms and organisations.

The degree course or disciplinary choice also determines the importance of IT skills. Chan and Fong (2018) developed a questionnaire for Hong Kong university students and revealed that compared to other generic skills, engineering students did not consider IT skills to be as important, however, they believe they are more proficient with IT skills than business students. This is expected as IT skills are a fundamental aspect towards the academic requirements (e.g., skills for programming and building hardware and software) for engineering students (Chan, Zhao, and Luk 2017). With regards to a business degree, IT skills are in general significantly less technical and usually require using software and programs to perform business-related work. For example, accountants are expected to use programs such as Excel, data analysis programs and cloud applications which are on average far less sophisticated than engineering programs such as MATLAB . As a result, this demonstrates that depending on the degree course, IT skills and other generic skills are usually seen as being more relevant to one discipline than to another (Chan and Fong, 2018).

4.3. Data

4.3.1. The Skills and Employment Survey (SES)

Alan Felstead and his colleagues developed the SES questionnaire to provide survey data on the skills and employment experiences of those working in the UK[17]. The sample representative of the dataset includes individuals that are between the ages of 20 and 65 year's old[18]. The sample period is between the years of 2012 and 2017. This has allowed an investigation of a wide range of generic job skills including computing skills which is a relatively new skill indicator in comparison with the other skills. Within the SES, the computing skills are measured between values of 0 to 4 whereby the former denotes an individual importance of computer use as "Does not apply in the job" and the latter denotes an individual importance of computer use as "Essential".

The SES consists of an additional explanatory variable which is not included in the Skills Survey. For example, the SES dataset provides considerably greater information on the use of computers in jobs such as "individual importance of computer use" which is not included in the Skills Survey. Another benefit of using the SES considers the estimation of computer uses in specific occupations as well as the level of intensity. Therefore, this provides a much more extensive understanding of how the wage structure is affected by modern technological advancements in different regions of the UK. One would hypothesise that the greater the computer skill an individual possesses, the greater earnings the individual would obtain (further explained in the data descriptive section 4.3.2.).

4.3.2. Descriptive statistics

Assessing the importance of generic skills proves challenging due to the diverse tasks encompassed across various occupations. Each occupation demands a distinct set of skills, and those deemed essential in one may not be relevant in another. Table 4.1 offers a detailed breakdown under each domain heading of each primary generic skill which are derived from

[17] As in 2006 and 2012, SES2017 comprised two different modes of interviewing: CAPI (computer-assisted personal interviewing: by interviewers) and CASI (computer-assisted self-interviewing: by respondents)

[18] Individuals included in the dataset also require having a paid job at which they work for at least 1 hour a week.

the SES. The objective is to examine the subheadings and gauge their significance in occupations, shedding light on which tasks fall under the categories of routine or non-routine.

The sub-heading “Dealing with people” is the most commonly considered essential task by a substantial margin. There are 72.06% of responders reported “Dealing with people” as essential in 2012 and 69.96% in 2017, whereas the next highest is "Importance of computer use" with 51.97% and 54.99% respectively. The essentiality of “Computer use” in 2017 has increased by 5.81% compared to 2012 which is expected as computers are exceedingly prevalent in occupations due to the advancement in technology. As such, researchers such as Brynjolfsson and Hitt (2000) as well as Bresnahan et al (2002) expressed the nuanced theory of the technical change in skill-biased where computerisation has rapidly substituted tasks which were previously labour intensive. Therefore, the use of computers can be argued as a routine task, although in some circumstances it also complements a non-routine task for higher educated interactive tasks e.g., programming.

In contrast, “Making speeches or presentations” represents the least essential task, recording values of only 11.78% in 2012 and 12.28% in 2017. Respondents reported that “More advanced mathematical or statistical procedures” are also less commonly regarded as essential, closely followed by “Planning the activities of others”. Notably, "More advanced mathematical or statistical procedures" saw a notable decline of nearly 17% between 2012 and 2017. This decline may be attributed to the integration of technology, particularly the growth of “Computer use,” which has replaced the need for advanced mathematical or statistical procedures in previously essential occupations. Thus, these three tasks are categorized as non-routine, as they are not regularly performed across most occupations. These findings echo the stability observed in task evolution in Britain, as highlighted by Francis Green (2012)."

From Table 4.1, it is evident that there are many occupations which require a wide range of literacy skills. In circumstances where employees work in teams or require cooperating with colleagues, it is imperative they possess the skill to exchange information through communication. As a consequence, a spectrum of literacy skills is required since writing and reading a crucial method to communicate with employees.

Table 4.1

The propensity of tasks deemed essential in Britain, 2012 to 2017

Tasks By Domain Heading	% task "essential"		Change % task "essential"
	2012	2017	2012-2017
Literacy			
Reading written information, e.g., forms, notices, or signs	44.34	43.59	-1.69
Reading short documents e.g., letters, or memos	41.47	40.96	-1.23
Reading long documents e.g., long reports, manuals, etc	28.19	28.83	2.27
Writing material such as forms, notices or signs	29.31	26.95	-8.05
Writing short documents, e.g., letters or memos	31.34	31.94	1.91
Writing long documents with correct spelling/grammar	22.35	23.44	4.88
Numeracy			
Adding, subtracting, multiplying or dividing numbers	32.53	29.31	-9.90
Calculations using decimals, percentages or fractions?	24.88	24.29	-2.37
More advanced mathematical or statistical procedures	15.30	12.70	-16.99
Communication: External			
Knowledge of products or services	40.47	41.44	2.40
Specialist knowledge or understanding	50.22	51.78	3.11
Knowledge of how organisation works	36.91	41.50	12.44
Selling a product or service	22.31	19.81	-11.21
Counselling, advising, or caring for customers or clients	39.50	35.36	-10.48
Dealing with people	72.06	69.96	-2.91
Communication: Influencing Others			
Instructing, training, or teaching people	33.66	31.37	-6.80
Persuading or influencing others	21.06	21.81	3.56
Making speeches or presentations	11.78	12.28	4.24
Planning the activities of others	15.50	17.36	12.00
Listening carefully to colleagues	43.81	43.98	0.39
Handling feelings of others	32.97	32.06	-2.76
Problem-Solving			
Spotting problems or faults	42.91	40.05	-6.67
Working out the cause of problems or faults	34.41	35.27	2.50
Thinking of solutions to problems	36.56	39.44	7.88
Analysing complex problems in depth	25.22	27.16	7.69
Physical			
Physical strength e.g., carry, push, or pull heavy objects	15.84	17.21	8.65
Work for long periods on physical activities	17.81	20.90	17.35
Skill or accuracy in using your hands or fingers	23.56	23.41	-0.64
Use or operate tools, equipment or machinery	31.91	27.83	-12.79
Computing			
Importance of computer use	51.97	54.99	5.81

Notes: Tasks are presented under the generic skills headings. For each task, respondents were asked "In your job, how important is . . . [each task]," answering from a scale from: "essential," "very important," "fairly important," "not very important," "not at all important/does not apply.

Similarly, the basic generic numeracy skills such as “Adding, subtracting, multiplying or dividing numbers” are often considered as essential in occupations. According to Rosenberg, Heimler and Morote (2012), the most important basic employability skills required for job performance are basic literacy, numeracy, and communication. However, from Table 4.2, it is evident that the more advanced mathematical procedures, the less likely the skill is regarded as essential. For instance, around a third of respondents deemed basic generic numeracy skills as essential, whereas 25% considered calculations which require decimals, percentages, or fractions as essential and only 15% for more advanced mathematical or statistical procedures. Therefore, from the perspective of job requirements, individuals are likely to have a basic understanding of numeracy skills, although many would not be required to practice more advanced numeracy skills.

Interestingly, all three domain headings for numeracy have decreased from 2012 to 2017 which suggests that "Numeracy" has become less essential in jobs over time. A possible explanation for the decrease in generic numeracy skills could be the progressive technological advancement in "Computer use" and the increased accessibility of technological equipment. Due to the growth of technological use in general, there could be an increased dependency on "Computer use" (5.81% from 2012 to 2017) which in turn mitigates individuals to use "Numeracy" skills as often as a result of the substitution from computers. This effect is exacerbated for "More advanced mathematical or statistical procedures" as computers are now able to calculate complex mathematical questions and can calculate faster and more efficiently than humans in most cases.

According to the National Association of Colleges and Employers (Koc and Koncz, 2009), employers consider communication skills to be the most important skill required and the most deficient skill for college graduates. From Table 4.1, there is a large discrepancy between the domain headings for the external and internal generic communication skills. In the field of work, on average it is seldom that individuals require the skill of "Making speeches or presentations". There is minimal opportunity for individuals to experience making speeches or presentations, thus a generic skill which is relatively redundant. Predominantly, speeches and presentations are conducted in a workplace with a greater number of employees tend to be presented by the supervisors or managerial positions. However, there are some communication skills that the respondents regarded as essential such as “Dealing with people” which supports the statement by the National Association of Colleges and Employers (Koc and Koncz, 2009).

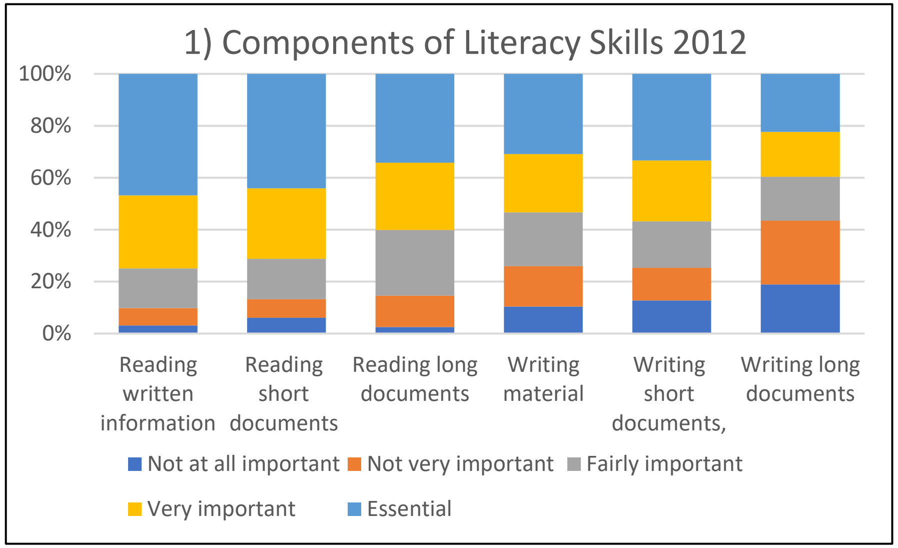
Finally, the generic skill “Physical” is lower in terms of essentiality to jobs compared to the other generic skills. This finding is consistent with Francis Green (2012) as they also reported a low rate of response that generic “Physical” skill is essential in occupations. In Table 4.1, “Physical strength” and “Work for long periods on physical activities” possess a low rate compared to other domain headings although they increased considerably between 2012 and 2017. On the other hand, “Use or operate tools, equipment or machinery” received a steep decrease between 2012 and 2017. Overall, this is expected as “Physical” skills are deemed to be less essential in occupations due to an increase in white-collar workers as opposed to blue-collar workers’ overtime.

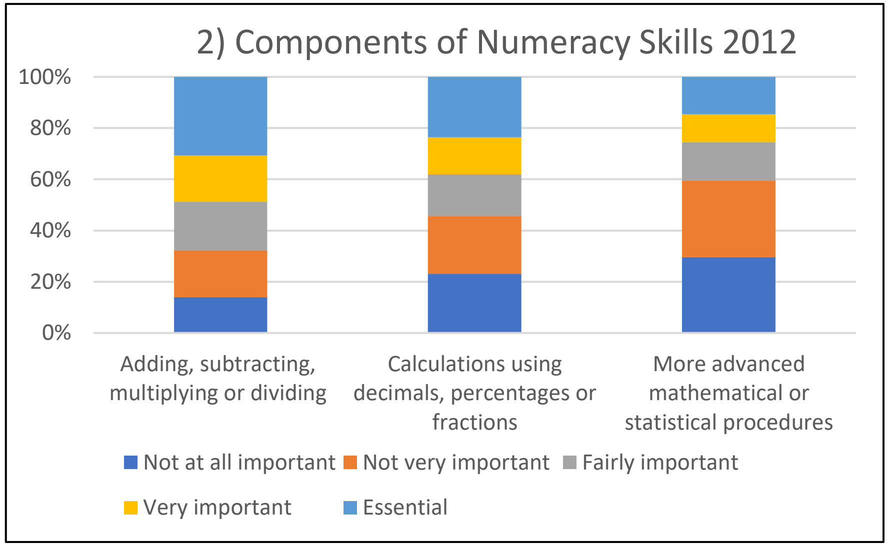
Table 4.1 already distinguished the percentages of respondents who classified each specific subskill to be essential for their occupation. Following on from this insight, Figure 4.1 and Figure 4.2 provide further depth as they illustrate the respondent's importance for each specific subskill of interest. constituents of the components of each generic skill for 2012 and 2017 respectively. The components of a subskill are divided into five separate constituents with a spectrum that spans from “not important at all” to “essential”.

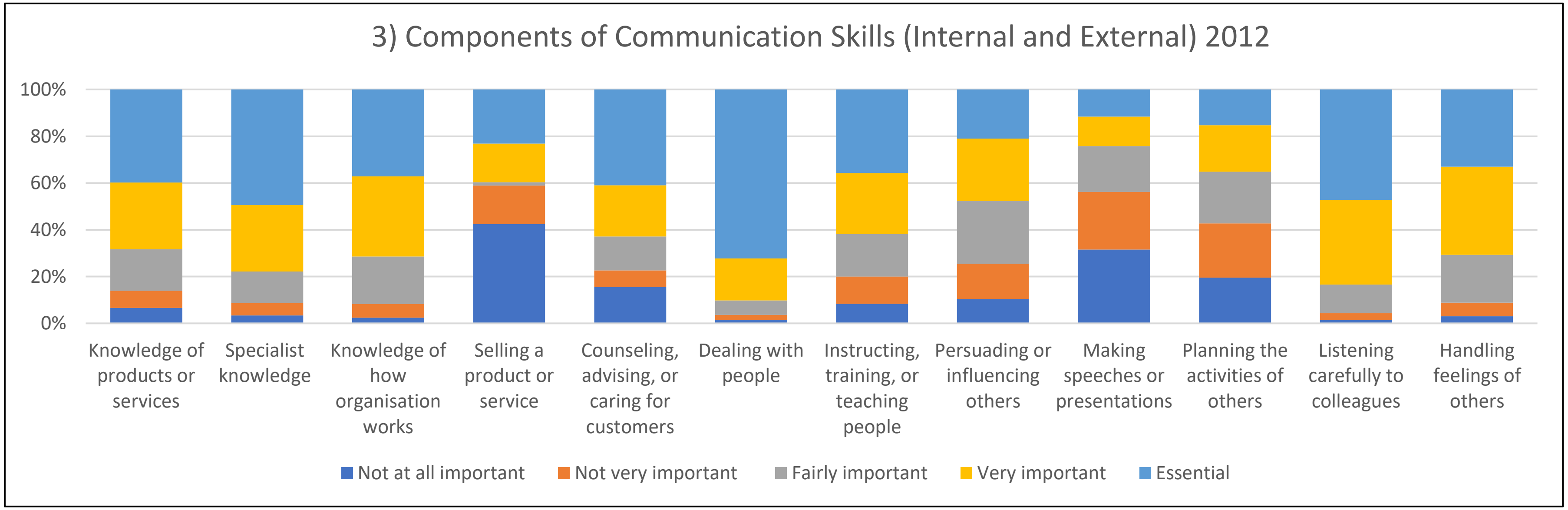
Figure 4.1.1 shows the components of literacy skills for 2012. There are not many jobs which do not include an element of literacy skills, which is translated as the components of literacy skills show a high rate of being “essential” to the respondent's occupation. The majority of literacy skill components exhibit a very low incidence of being deemed "not at all important," particularly notable in the cases of two subskills: “reading written information” and “reading long documents”. Among these components, “writing long documents” stands out with both the lowest percentage of being essential and the highest percentage of being considered “not at all important” as a job requirement.

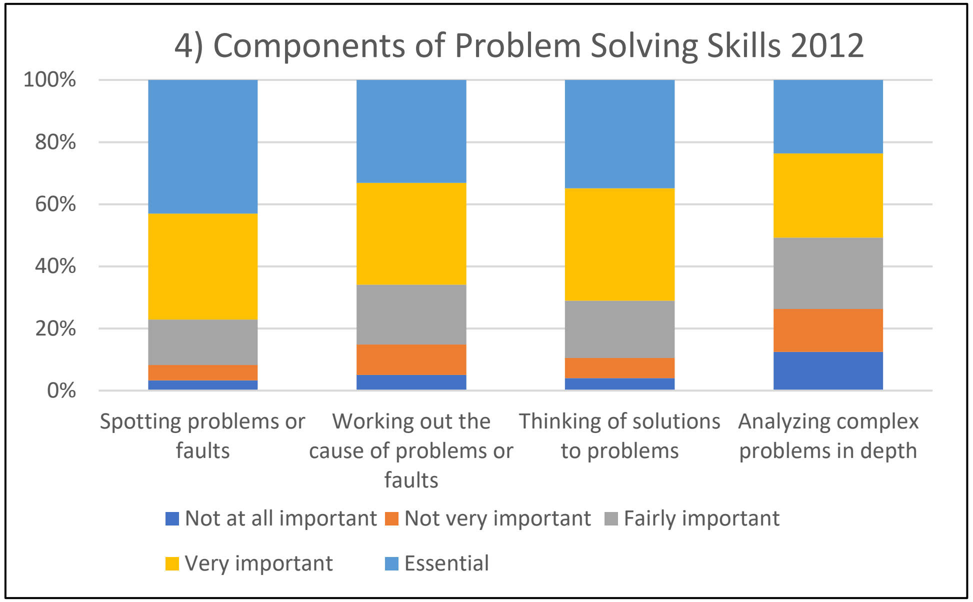
Figure 4.1

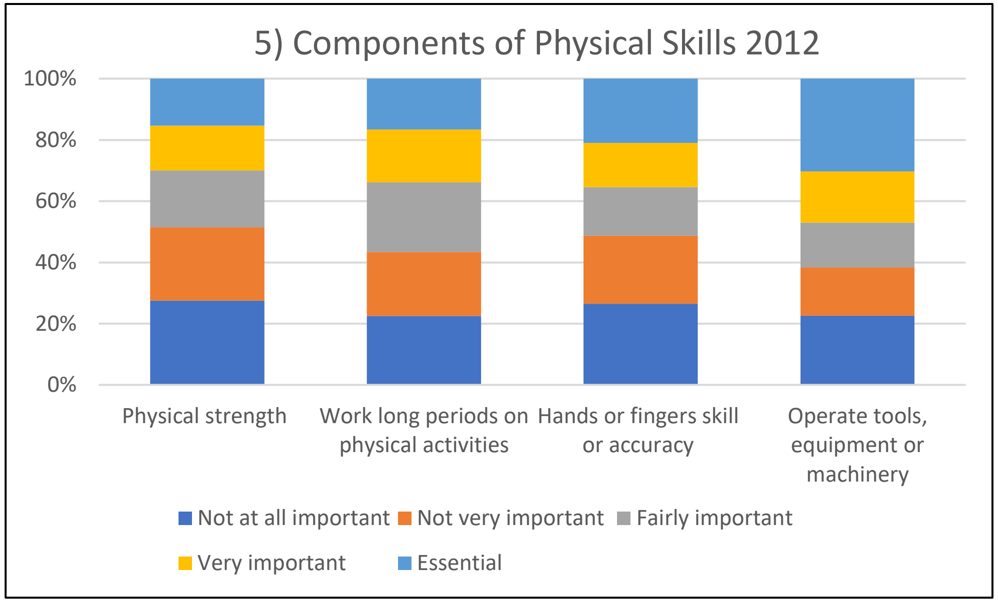
Constituents of the components of each generic skills (2012)

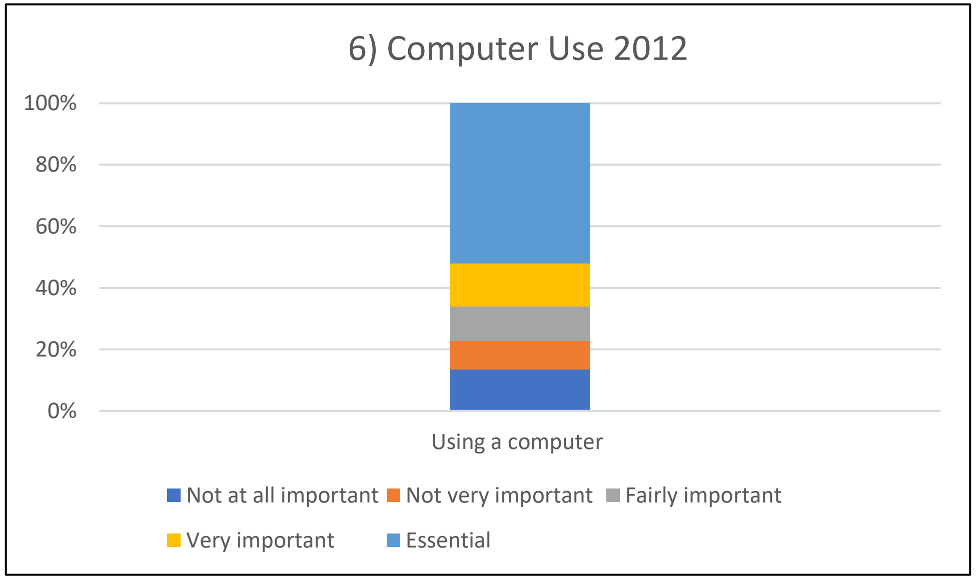


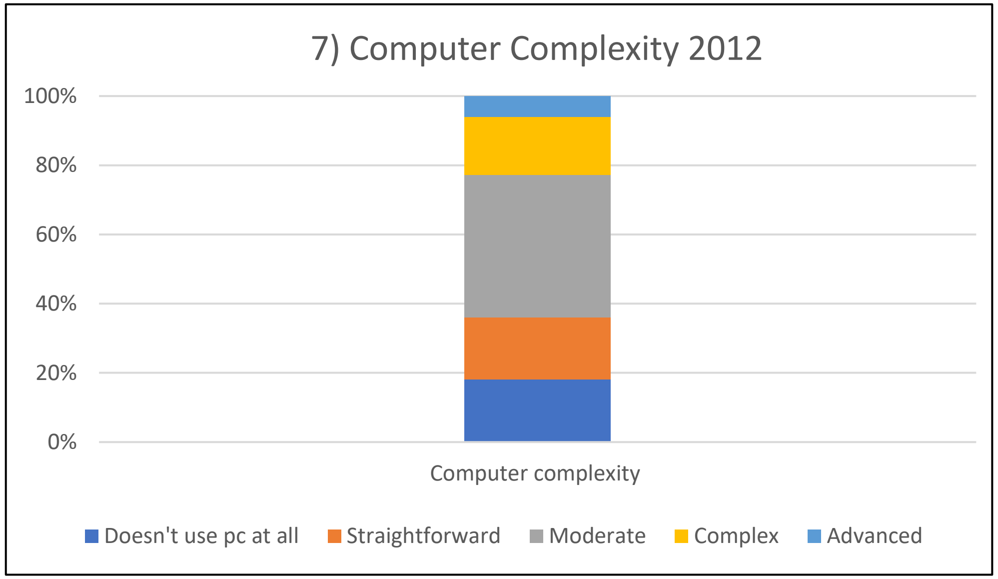








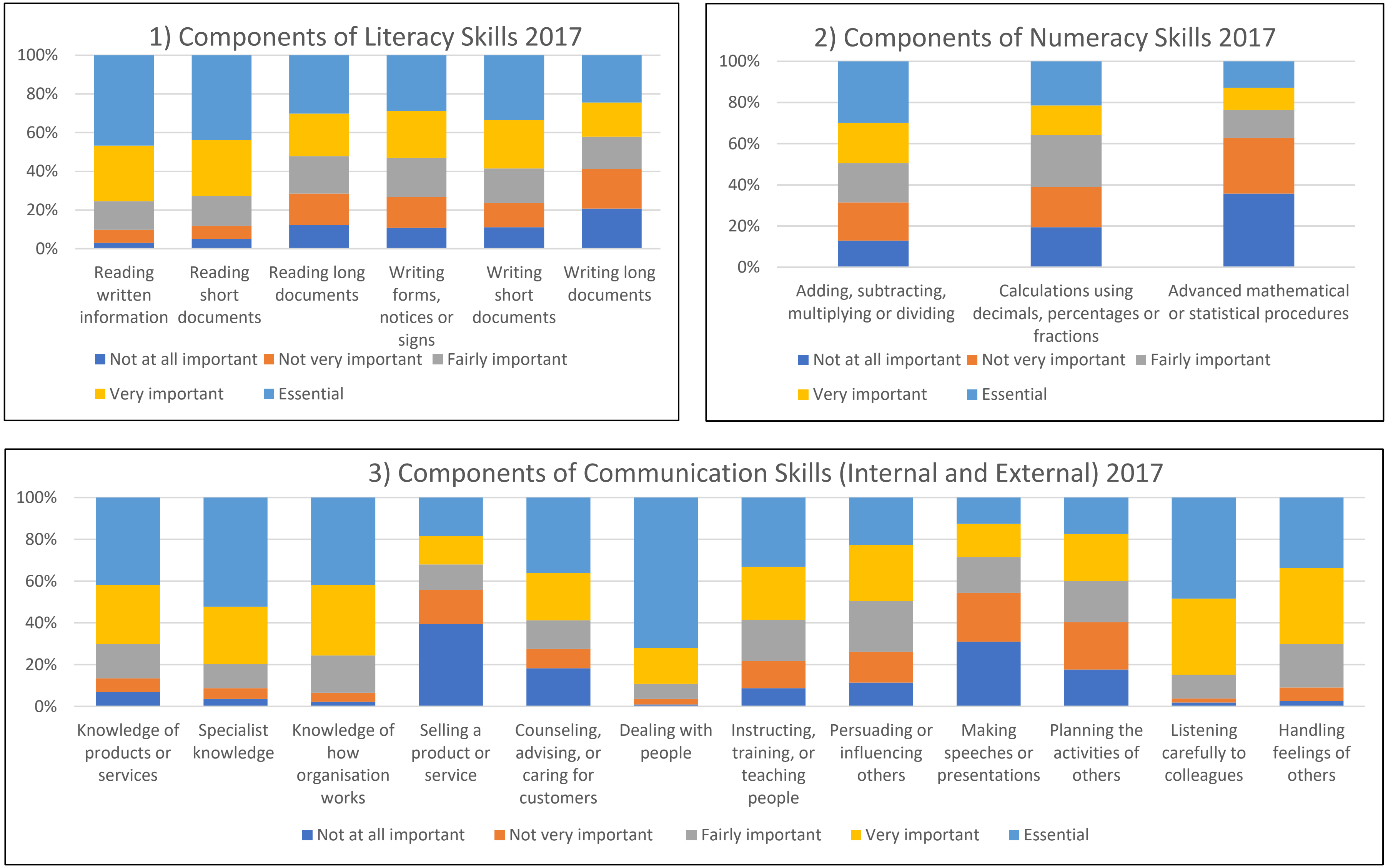


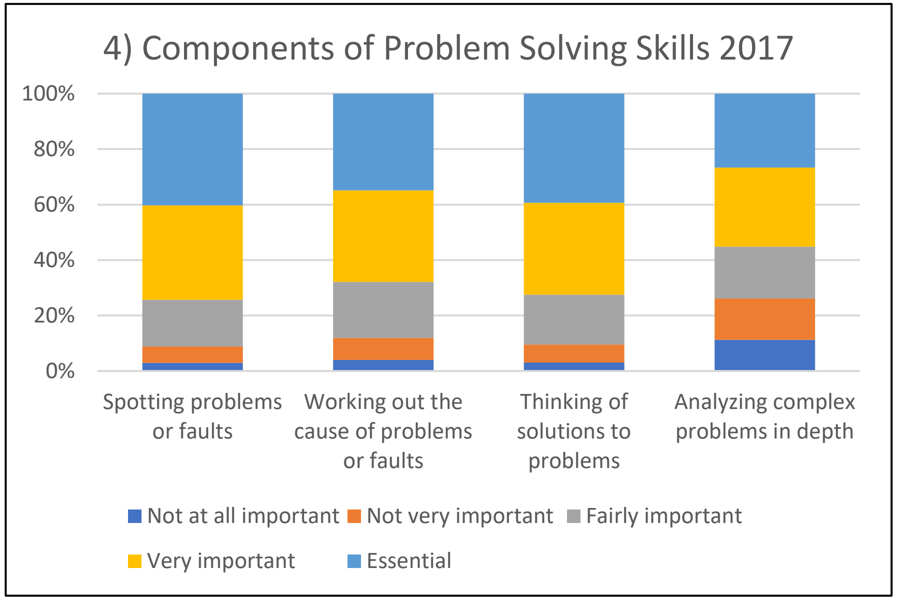


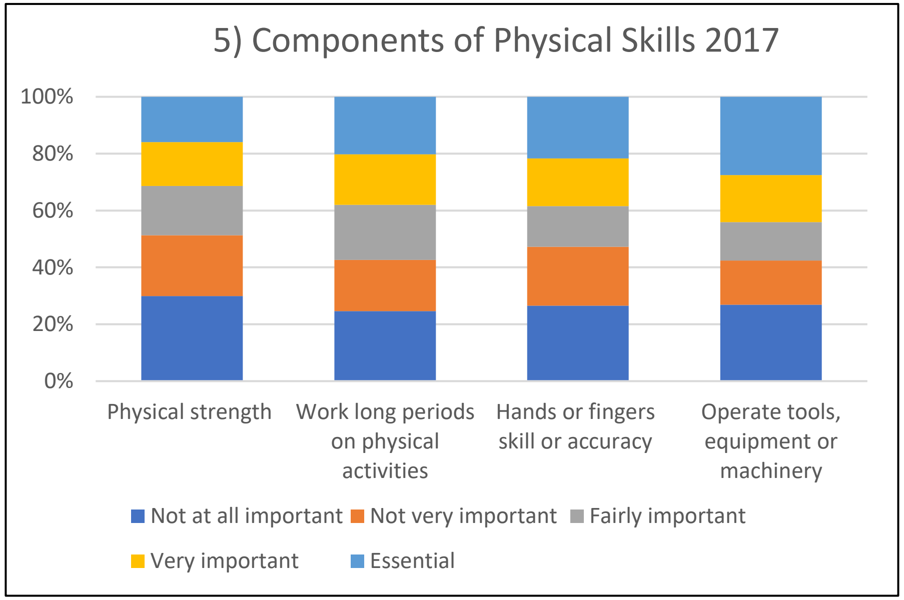
Note: The bar chart shows the constituent of importance as a percentage for each task. Computer complexity has a different measurement, with a scale of "Doesn't use pc at all", "Straightforward", "Moderate", "Complex" and "Advanced".

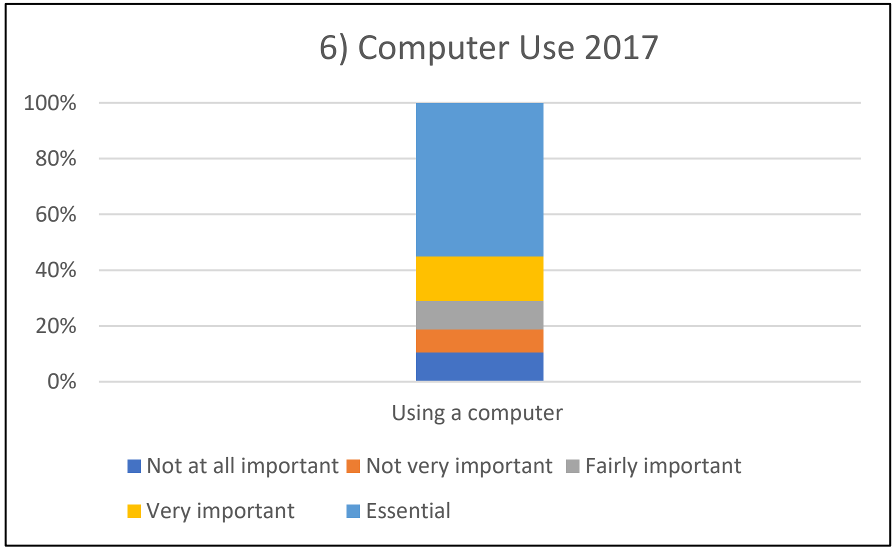
Figure 4.2

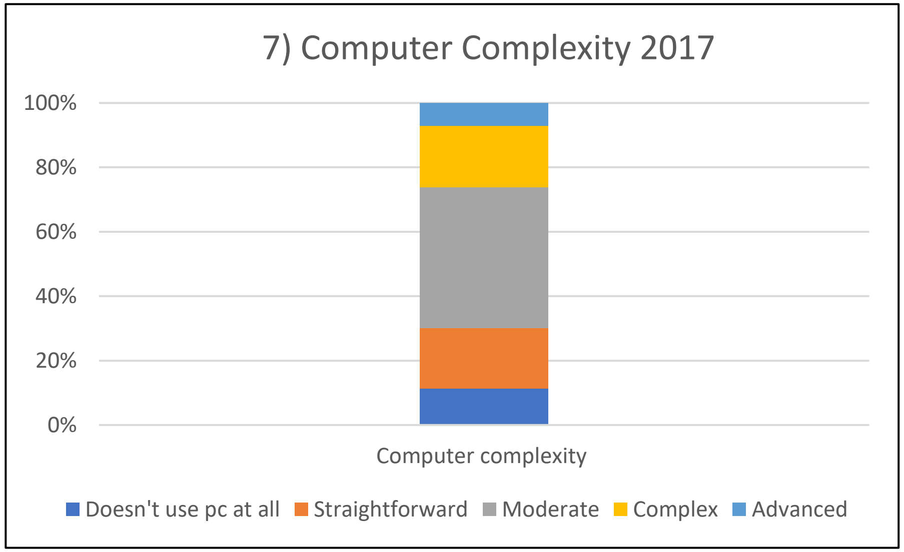
Constituents of the components of each generic skills (2017)











Note: The bar chart shows the constituent of importance as a percentage for each task. Computer complexity has a different measurement, with a scale of "Doesn't use pc at all", "Straightforward", "Moderate", "Complex" and "Advanced".

On average, the components of numeracy skills in 2012 exhibit a higher prevalence of being considered “not at all important” compared to literacy skills, as depicted in Figure 4.1.2. Specifically, over 30% of respondents express that “more advanced mathematical or statistical procedures” are perceived as “not important at all”. This outcome is anticipated given the limited number of occupations that necessitate advanced mathematical or statistical knowledge, especially considering the increase in technological utilisation and advancements. Nevertheless, basic mathematical skills are perceived as relatively important, with over 30% of respondents categorizing "adding, subtracting, multiplying, or dividing" as "essential.

Half of the communication skills components exhibit a notably low incidence of being considered “not at all important” as shown in Figure 4.1.3. Conversely, these subskills demonstrate a relatively high occurrence of being deemed "essential" in various occupations. Among these, the subskill “dealing with people” stands out the most, with the lowest rate of being considered “not at all important” and the highest rate of over 70% of respondents perceiving the skill as “essential”. Furthermore, across many components, a substantial proportion is labelled as “very important”. For example, approximately 30% of respondents consider "handling the feelings of others" as “essential”. When incorporating the category of “very important”, this figure increases to over 70% for this particular subskill. This implies that it is beneficial to consider other constituents of the importance of a skill.

In general, the “problem-solving” skills components in 2012 were widely regarded as “essential” as seen in Figure 4.1.4. This is especially true for the subskill “spotting problems or faults” since almost 80% of respondents consider this skill as “very important” or “essential”, this is closely followed by “thinking of solutions to problems”. As skills become more advanced or complex, there is a decrease in the proportion of respondents labelling the skill as “essential” and an increase in those deeming it “not at all important”. This trend is evident in the assessment of “analysing complex problems in depth”, which mirrors the patterns observed in advanced mathematics and statistical procedures.

Figure 4.1.5 shows that the components of “physical” skills in 2012 have a relatively high ratio of categorised as “not at all important”. Over 20% of the constituents within each of the four subskills are labelled as “not at all important”. Compared to the other generic skills components, physical skills exhibit a higher percentage of either “not at all important” or “not very important”. This is expected given the significant shifts in the foundational restructuring of the UK economy over recent decades, which is primarily driven by changes in technology. The Centre for Vocational Education Research (CVER) discovered a decline in manual skills across all employment sectors, while the use of analytical and interpersonal skills has increased significantly.

Figure 4.1.6 and 4.1.7 illustrates the “computer use” and “computer complexity” for 2012 respectively. It is important to highlight that in the case of “computer complexity”, the metrics represent the complexity of the computer rather than the level of usage. At first glance, it is apparent that more than half of the respondents consider computer use as “essential” in their respective occupations. Whereas only 6% of respondents indicate the need to possess an “advanced” level of computer complexity within their occupation. There is almost 80% of respondent’s state using “computer complexity” at levels ranging from “doesn't use pc at all”, “straightforward” or “moderate”. Therefore, while the majority of respondents use computers, few use computers at a “complex” or “advanced” level.

Figure 4.2 illustrates the constituents of the components of each generic skill in the year 2017. The results between 2012 and 2017 produce relatively similar results, although there are a few notable points. For example, the only noticeable difference in the components of “literacy” skills between the two periods is “reading long documents”. There was a significant increase in the subskill during 2017 as seen in figure 4.2.1. This is likely due to the increase in digitalisation and greater transition to media use.

Additionally, in 2017, the elements of other general skills exhibited no significant deviation from the components observed in 2012. Notable variations were only observed in numeracy and computer skills. Specifically, the subskill related to “calculations using decimals, percentages, or fractions” demonstrated an increase in responses categorized as “fairly important”, as depicted in Figure 4.2.2. Furthermore, the responses indicating “not at all important” declined for all numeracy skill components, except for “more advanced mathematical or statistical procedures. As anticipated, the subskill during 2017 experienced an increase in responses categorized as “not at all important”, reflecting the growing influence of technology and its advancements between 2012 and 2017. This shift indicates that technology has substituted the demand for advanced mathematical or statistical expertise in occupations.

Figure 4.2.6 and 4.2.7 demonstrates that in 2017, there was a rise in both the utilisation and the “computer complexity” respectively. “Computer use” experiences a decline in the “not at all important” and an increase in the “essential” response. Similarly, “computer complexity” experienced a decrease in responses indicating “doesn't use a PC at all” and an increase in those denoting “complex” and “advanced” in 2017 compared to 2012. Therefore, the computer components and numeracy components are negatively correlated, especially for the more advanced mathematical and statistical procedures, as previously explained.

Table 4.2 reports the coefficients from the estimates of the wage equation when the 2012 and 2017 waves of the SES are pooled together. The generic skills are computed from the PCA statistical procedure which implements a dimension reduction method. This transforms the large set of related subskill variables into one which represents a specific generic skill while able to retain the overall information. The default is to choose the first eigenvalue when estimating the influence of the generic skill since the first eigenvalue accounts for the most variance. The eigenvector associated with the largest eigenvalue determines the most stable distribution of the specific generic skill of interest given the constitutes of the generic skills subset questions. Other control variables such as ethnicity are included to determine whether there is a wage penalty experienced by ethnic minorities. Due to the lack of sample size for the ethnic minority, the regression estimates the comparison between White and Non-White ethnic groups. The dependent variable is the log of real hourly wages.

There are three rotation techniques examined which includes the unrotated PCA, as well as two rotated methods: Promax and Varimax. Promax is an oblique rotational method, allowing for correlations between factors, while Varimax is an orthogonal rotational method, emphasizing independence between factors. Rotation is essential to capture the maximum variance of each generic skill. If the components remain unrotated, the efficacy of PCA diminishes, necessitating a higher number of components to elucidate the variance of the generic skill.

Table 4.2

Regression of wage equation with PCA: Generic skills and controls

Variables	Unrotated	Promax	Varimax
Female	-0.124*** (0.011)	-0.111*** (0.011)	-0.110*** (0.011)
Age	0.031*** (0.003)	0.028*** (0.003)	0.028*** (0.003)
Age Squared (x1000)	0.0301*** (0.003)	0.0303*** (0.003)	0.0302*** (0.003)
Tenure	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Tenure Squared^{x100}	0.000** (0.000)	0.000** (0.000)	0.000* (0.000)
Ethnicity			
White	-	-	-
Non-White	-0.028*** (0.018)	-0.040** (0.018)	-0.042** (0.018)
Generic Skills			
Literacy Skills	0.018*** (0.004)	0.039*** (0.006)	0.037*** (0.006)
Numerical Skills	-0.001 (0.004)	-0.013** (0.006)	0.001 (0.006)
Communication Skills	-0.001 (0.003)	0.049*** (0.006)	0.051*** (0.006)
Problem-Solving Skills	0.031*** (0.004)	0.023*** (0.005)	0.012** (0.006)
Physical Skills	-0.059*** (0.004)	-0.081*** (0.006)	-0.082*** (0.006)
Computer Complexity			

Doesn't use PC at all	-	-	-
Straightforward	0.038** (0.018)	0.045** (0.018)	0.037** (0.018)
Moderate	0.133*** (0.018)	0.150*** (0.017)	0.139*** (0.017)
Complex	0.183*** (0.021)	0.197*** (0.021)	0.186*** (0.021)
Advanced	0.250*** (0.027)	0.260*** (0.027)	0.247*** (0.027)
Constant	**1.725*** (0.082)**	**1.750*** (0.082)**	**1.753*** (0.083)**
Observations	**4479**	**4479**	**4479**
R-squared	**0.5109**	**0.5116**	**0.5103**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places. The base category is the "White British" group and the "Doesn't use PC at all"

The first row of results from Table 4.2 shows the estimates of the females which is represented with a simple (1,0) dummy variable. There are three separate figures which illustrate the results of the unrotated, promax and varimax estimates. It is evident from all three rotational techniques that, on average, females earn less than their male counterparts, this is statistically significant at the 1% level. The rotated outcomes for females exhibit remarkable similarity, with coefficients of -0.111 and -0.110 for promax and varimax, respectively. In contrast, the unrotated result deviates slightly with a figure of -0.124. Given the retention of more than one factor, rotation becomes essential; otherwise, the interpretation of results becomes challenging. Without rotation, the impact of PCA diminishes, necessitating a greater number of components to explain the variance in the dataset.

As a result, focus will be placed on the outcomes generated by rotational techniques. Upon comparing the second and third columns, it is evident that the results for generic skill variables and other control variables are very similar. Therefore, the selection of the rotational method does not significantly impact the results. However, considering the high correlation among most of the generic skills as demonstrated by the Pearson correlation in Table 4.3, the Oblique rotational technique will be adopted, as detailed in the methodology.

Table 4.2 is intended to provide insight into the results of a baseline estimation by pooling the sample in a wage equation regression. Further detailed estimations will be performed in the results section. In Table 4.2, the generic skills that demonstrate a statistically significant (at a 5% level or better) and positive impact on wages are literacy skills, communication skills and problem-solving skills, with coefficients of 0.039, 0.049 and 0.023, respectively. Given that the measure of the first five generic skills (excluding computer complexity) has a unit variance, and the dependent variable is the log of real wages, these coefficients indicate the proportional increase in wages for a one standard deviation increase in each skill measure. This implies that an individual with literacy skills (communication | problem-solving) one standard deviation above the mean experiences a 3.9 (5.0 | 2.3)% higher wage than an individual with the mean level of literacy (communication | problem-solving) skills in the sample.

The generic skills correlated with lower wages include numerical skills and physical skills, with coefficients of -0.013 and -0.081, respectively. Initially, it may seem surprising that an increase in the importance of numerical skill results in a negative wage premium. However, this can be largely attributed to the strong correlation with computer complexity, which mitigates the positive impact of numerical skills. This perplexity will be further discussed in section 4.5.

The latter section of Table 4.2 presents the wage premium associated with different levels of computer complexity required in an occupation. The five categories of computer complexity range from “doesn't use PC at all” to “advanced”. The coefficient for each category is based on “doesn't use PC at all” as the baseline. As anticipated, there is a notable increase in the wage premium as the complexity of computer use rises. All coefficients are statistically significant, and there is a substantial increase in wage premium with each increment in computer complexity. For example, the “computer complexity” transitioning from “straightforward” to “moderate”, the wage premium increased from the baseline by 4.5% to 15%. Notably, computer complexity exhibits the most substantial increase in wage premium compared to other generic skills. An individual requiring an “advanced” level of “computer complexity” in their job on average earns 26 percent more than an individual who “doesn't use PC at all”.

4.4. Methodology

4.4.1. Types of skills

To assess the benefits of generic skills, it is essential to identify and define the various types of the chosen generic skills. In this research, there are six generic skills which are investigated and are shown as follows:

1) **Literacy Skills:** both reading and writing forms, notices, memos, signs, letters, short and long documents, etc.
2) **Numeracy Skills:** adding, subtracting, division, decimal point, or fraction calculations, etc., and/or more advanced maths
3) **Communication Skills:** a range of related managerial skills, including persuading or influencing others, making speeches or presentations, writing long reports, analysing complex problems in depth, dealing with people, selling a product or service, counselling, or caring for custom
4) **Problem-Solving Skills:** detecting, diagnosing, analysing, and resolving problems
5) **Physical Skills:** the use of physical strength and/or stamina
6) **Computer Complexity:** use computers effectively, including related equipment and computer software i.e. computer literacy

These six essential skill components were selected based on various considerations. Firstly, retaining this number of factors resulted in skill types that are easily interpretable. Secondly, an initial descriptive PCA test revealed that these six components accounted for more than 75% of the variability in the activities. The base category skill components were closely associated with the already chosen generic skills, adding minimal extra information that could impact wage returns significantly. For instance, technical know-how skill showed a high correlation with physical skills. Nevertheless, even after factor rotation, the six chosen generic skills remain distinct and identifiable.

4.4.2. Applying PCA to determine the generic skill indices

From the SES questionnaire, 30 subsets of relevant skill variables have been extrapolated to depict the significance of an individual's particular occupation. To effectively estimate the wage equation, the implementation of a data reduction procedure is necessary. The Principal Component Analysis technique will be used to reduce the data and address the issue of multi-collinearity. In the case of computer skills, a different approach was taken. Instead of measuring computer importance, the complexity level of computer use was chosen.

The PCA's objective is to determine satisfactory constrained variables from underlying unobserved components that encompass a substantial group of observed variables. However, a notable limitation of this method of this method is the absence of a specific principle to determine the optimal number of components to extrapolate. The chosen number of components relies on the combination of the underlying data as well as the theoretical basis such as the interpretability of the components. The primary aim of the PCA in this paper is to identify the inherent component of the generic skills by deriving indices from a sizeable group of skill components that can be applied in wage estimation.

To prepare the PCA, it is common practice to assume linearity. Therefore, for the PCA composition, the ordinal degree of "importance" for each specific generic skill variable is translated into an increasing cardinal scale, ranging from a value of one (not at all important) to five (essential). To assess the relevance of employing a data reduction process, an examination of the correlation matrix of the group of variables for each specific generic skill is necessary. While the matrix is not provided here for the mean of brevity, the matrix indicates high correlations among variables. Furthermore, performing the overall Kaiser-Meyer-Olkin measurement of sampling adequacy generates a coefficient above 0.9. In terms of the individual Kaiser-Meyer-Olkin measure, most of the variables surpass a coefficient of 0.9, with only four variables with a value below 0.8. As a result, this is strong evidence to suggest that the sample data is appropriate to undertake the PCA procedure.

The following stage of the PCA is to determine the optimal rotation of the data points in Euclidean space such that the variance is maximised along the first axis (Greenacre et al., 2022). The objective of rotation is to improve the interpretability of the extracted factors and comprehension of the primary components. In turn, a structure of the loading matrix called “simple structure” is obtained, representing a cluster of correlated variables. There are two primary forms of rotations.

This first type of rotation is the Orthogonal rotation. This technique preserves the orthogonality by shifting the factors within the factor space while retaining 90-degree angles between them to attain optimal simple structure. This rotation method ensures that the factors remain zero correlation relationship as the cosine of the angles between unit length vectors are exactly zero for a 90-degree angle. As a result, this rotation can enhance the interpretability of the results due to the factors being uncorrelated.

The alternative rotation is the Oblique rotation. The Oblique rotation approach enables correlations between the factors. The angles between the factors can either exceed or fall below the 90-degree angle. The most popular and commonly used oblique rotation method is Promax. The technique Promax is the most commonly used as it delivers effective solutions consistently with a tendency to yield reproducible results compared to other oblique rotational methods.

To determine the technique for factor rotation implementation, an examination of the correlation matrix of the generic skills is necessary. Tabachnick and Fiddell (2007) suggest checking if the factor correlation matrix is around 0.32 and above. In the circumstance of correlation surpassing 0.32, it indicates a 10 percent or greater overlap in variance of factors which in turn is a sound justification for using oblique rotation (in this case, the Promax technique).

Table 4.3 illustrates the factor correlation matrix between the generic skills by utilising Pearson's correlation method in STATA. Pearson's matrix shows that the majority of the correlations between the skills are above 0.32. This is similar to previous papers such as Dickerson and Green (2004), the non-manual generic skills are highly correlated with each other. For instance, literacy skills have a higher correlation than 0.32 when matched with other skills except for physical skills. It is expected that physical skills are weakly or even negatively correlated with other skills since the labour market does not value physical work as highly as other skills which require further education or qualifications to attain. Physical skills are only positively correlated to communication and problem-solving skills, though the degree of correlation is weak. Moreover, communication skill has a relatively low correlation with the other skills. Apart from the association with literacy skills, the correlation coefficients for all other general skills are below 0.32 when compared to communication skills.

Numeracy skills are strongly linked with computer skills since to excel at software programming, individuals need to be knowledgeable at writing good algorithms and computer tasks (Arenas, Paragulla and Beltran, 2019). This finding is consistent with the obtained result since numeracy skills are highly correlated with both computer skills and computer complexity. Except for physical skills and literacy skills, computer complexity is highly correlated with non-manual generic skills which indicates that most skills require a strong association of literacy skills. This could be argued that literacy could be the most important generic skill as to develop other skills, one would likely require a good foundation of literacy skills.

Overall, given that the majority of correlation coefficients are above 0.32, the most suitable option is oblique rotation, specifically the Promax rotation. The Promax method is the most parsimonious simple structure as the factors are correlated. There are three main steps to Promax rotation. Initially, the factors are rotated orthogonally. Followed by generating a target matrix by elevating the factor structure coefficients to an exponent greater than two. Finally, the matrix experiences a "Procrustean" rotation for a best-fit outcome. Promax is often the preferred oblique rotation method since it is relatively simple to exercise, generally delivers effective solutions and yields more reproducible results compared to alternative techniques such as direct Oblimin rotation.

An advantage of using an oblique rotation strategy is the ability to align the solution more closely to the researcher's perspective on the objective honours. Also, since there are more parameters estimated in an oblique rotation, this typically provides a more accurate fit to the sample data compared to orthogonal solutions (Kieffer, 1998). However, some of this fit can incorporate overfitting. This occurs when a model is too complex and fits the sampling data too closely, which causes sampling error variance. As a result, despite orthogonal solutions often fitting sample data adversely, generally, the strategy exhibits higher replicability in the future sample as orthogonal solutions require less sampling error.

Table 4.3

Correlations between generic skills and computing skills

	Literacy Skills	Numeracy Skills	Physical Skills	Communication Skills	Problem-Solving	Computer Complexity
Literacy Skills	1					
Numeracy Skills	0.4346	1				
Physical Skills	-0.076	-0.0574	1			
Communication Skills	0.3753	0.2473	0.0756	1		
Problem-Solving	0.5208	0.3912	0.1315	0.3211	1	
Computer Complexity	0.4155	0.4673	-0.3272	0.2036	0.3567	1

Notes: For the 5 continuous generic skills and computer complexity, the correlations reported are Pearson correlation coefficients.
Generic skills (row 1 to 5) are measured as an index running from 1 ("not at all important") to 5 ("essential") and hence the correlations reported are Spearman rank order correlation coefficients. Computing complexity (row 6) is measured as an index running from 0 ("don't use PC at all") to 4 ("advanced") and hence the correlations reported are Speakman rank correlation coefficients.
All the correlation coefficients in the Table are significantly different from zero at the 1% level.

4.4.3. Model setup and parameters of interest

4.4.3.1 Eigenvector and eigenvalue of the PCA

Once the optimal PCA technique is performed, it generates the eigenvectors of the covariance matrix of that dataset which represents the directions of maximum variance in the data. Each eigenvector corresponds to a principal component, and they are arranged in order of their associated eigenvalues, indicating the amount of variance explained by each component. The initial principal component accounts for the highest variance, followed by subsequent components.

In the results, the first principal component is chosen since it generates the highest eigenvalue of a matrix. This highest eigenvalue often reflects the dominant behaviour or characteristic of the system characterised by the matrix. For instance, in the context of generic skills, the highest eigenvalue might signify the primary scaling factor, signifying the most influential subset question. A larger influence corresponds to greater weighting. Additionally, the highest eigenvalue can signify the stability of the system. When its magnitude exceeds one, it suggests that the corresponding principal component captures more variance than a single original variable. This is typically advantageous as it denotes that the principal component is informative and significantly aids in dimensionality reduction.

In summary, choosing the highest eigenvalue of a matrix can be crucial in various mathematical analyses, stability assessments, optimization problems, and data analysis tasks. It often provides key insights into the dominant behaviour or characteristics of the system described by the matrix.

4.4.3.2 Quantile Regression

When the PCA and rotational strategy is performed, the sample is estimated using quantile regression. The quantile regression (QR) framework offers a practical method for understanding how covariates affect an outcome at different points along its distribution. Using the empirical study as an example, when examining the impact of various generic skills on the wage premium, the QR framework can identify the generic skill with the greatest comparative advantage for a specific gender and ethnic group. This technique enables the determination of the impact of a covariate for individuals precisely at the 10th percentile of the wage distribution and thus can compare with the impact experienced for individuals at the 90th percentile. As a result, this provides the ability to comprehend how individuals are affected by a covariate at any position of the wage distribution spectrum.

Most QR empirical paper utilises conditional quantile regression (CQR) since it examines the impact of a covariate on a quantile of the outcome conditional to the values of other covariates (Borah and Basu, 2013). However, there is a notable limitation of the CQR since the interpretation of the results is constrained when the effects for numerous conditional quantiles vary. Therefore, this causes challenges as the estimated results from a CQR do not effectively provide insight into policy circumstances that are associated with the relevant covariates. As a result, the implementation of the unconditional quantile regression (UQR) method will be employed to address the limitations imposed by the CQR framework.

In this study, the investigation focuses on how generic skills influence the wage premium effect of an individual across the earning distribution, as opposed to using OLS to determine the mean. For example, instead of focusing on the effect of a covariate on the mean wage, a more insightful measure of interest is how the covariate affects the lower (higher) quantile of the earnings distribution. Both CQR and UQR can perform this empirically, although CQR cannot average up to their unconditional population quantiles, the UQR is beneficial in this aspect.

4.4.3.3. Unconditional quantile regression

To understand the conceptualisation of UQR, it is important to comprehend the mechanics of influence functions (IF) and their relations to the unconditional distribution of the outcome. In this research context, assume an individual earning (y) as a random variable characterised by a cumulative distribution function F_p and a density distribution $f_p = dF_p$. The statistic of interest q can be expressed as a function of the cumulative distribution function (CDF):

$$q_{Fp} = q(F_p) \tag{4.1}$$

Now consider there is an additional labour in the model represented as the second distribution q_p^0, with earnings equal to y_0 given the original size N. With the additional labourer, the distribution could be written as:

$$C_p^0 = \frac{N}{N+1} F_p + \frac{N}{N+1} * \gamma(y \geq y_0) \tag{4.2}$$

With both the distributions in consideration, the alteration in the distributional statistic resulting from the additional labourer is the variance between q_{Fp} and $q_{C_p^0}$. By adjusting this variance by the relative change in the population size, the influence function is obtained. Let $\theta = (N+1)^{-1}$, then the IF of a labourer with earning y_i on the statistic q can be simplified as:

$$IF(y_i, q, F_p) = \lim_{\theta \to 0} \frac{q\{(1-\theta)F_p + \theta * \gamma(y \geq y_i)\} - q(F_p)}{\theta} \tag{4.3}$$

Equation (4.3) demonstrates the Gateaux derivative and underlies the first order linear approximation of how the earnings y_i of a labourer impact the distributional statistic C. Incorporating the IF into an UQR framework involves introducing the concept of the Recentered Influence Function proposed by Firpo et al (2009):

$$\mathrm{RIF}\{y_i, q, F_p\} = q(F_p) + IF(y_i, q, F_p) \tag{4.4}$$

To focus on the analysis of unconditional quantiles, Firpo et al (2009) define the RIF as:

$$\mathrm{RIF}\{y_i, U_\varphi(.), F_p\} = U_\varphi(p) + \frac{\varphi - \omega\,\{y_i \leq U_\varphi(p)\}}{f_p\{U_\varphi, (p)\}} \tag{4.5}$$

Where $U_\varphi(p)$ represents the unconditional quantile, with φ denoting specific percentile of interest. The variable ω serves as an indicator function to determine whether the observation y_i falls below $U_\varphi(p)$, and $f_p\{U_\varphi,(p)\}$ represents the density function of distribution of y evaluated at the $U_\varphi(p)$.

From equation (4.5), the estimation of the unconditional mean $\mathrm{RIF}\{y_i, U_\varphi(.), F_p\}$ is the process of the estimation of an unconditional probability model when an observation is situated below or above the quantile of interest. This is adjusted by a scaling factor to account for the significance of the quantile within the distribution and recentred by a constant.

The most common estimation of UQR is the RIF regression approach. This method can be used to estimate a RIF regression which encompasses a large set of distributional data, including the practice of quantile effects (Avila, 2020). However, the STATA command "uqreg" will be employed for empirical estimations, enabling the utilization of unconditional quantiles and binomial models for the UQR regressions.

After the rearrangements of equation (4.5), this can be illustrated as:

$$\mathrm{RIF}\{y_i, U_\varphi(.), F_p\} = U_\varphi(p) + \frac{p-1}{f_p\{U_\varphi,(p)\}} + \frac{1\{y > U_\varphi,(p)\}}{f_p\{U_\varphi,(p)\}} \tag{4.6}$$

Equation (4.6) indicates that once the distribution p is determined, obtained the estimation for the quantile $U_\varphi(p)$ and the density of the distribution at a specific quantile, $f_p\{U_\varphi,(p)\}$, then the binary variable, $1\{y > U_\varphi,(p)\}$ is the only parameter that requires modelling.

Therefore, when incorporating the RIF regression model into an OLS, analysis of the outcome involves using a linear probability model. However, one can utilise a binomial model for better model estimation by the use of Unconditional Partial effect (UPE) which applies the rescaled average marginal effects (Firpo et al, 2009). One example of an UPE is the command, uqreg which is used for the estimation.

Firpo et al (2009) demonstrated that when represent the conditional expectation of $\delta \mathrm{RIF}\{y_i, U_\varphi(.), F_p\}$ as a function of explanatory variables (including the control variables e.g. gender, age, marital status etc.), $E\left[\mathrm{RIF}\{y_i, U_\varphi(.), F_p\} \middle| \mathrm{Z} = \mathrm{z}\right] = M_\varphi(\mathrm{z})$, the RIF regression is considered as an UQR. Since $E_Z E\left[\mathrm{RIF}\{y_i, U_\varphi(.), F_p\} \middle| \mathrm{z}\right] = U_\varphi(.), F_p$ by the RIF definition, $E_Z\left[\frac{\delta M_\varphi(z)}{\delta Z}\right]$ refers to the marginal effect of a minimal shift in the distribution of covariates on the φth unconditional quantile of y_i, assuming ceteris paribus holds.

Now identify the linear function used to model $1\{y > U_\varphi, (p)\}$ as $M(\mathbf{X}\mu)$, so the UPE shown as equation (4.6) becomes:

$$E\left[\frac{\delta \mathrm{RIF}\{y_i, U_\varphi(.), F_p\}}{\delta X}\right] = \frac{1}{f_p\{U_\varphi, (p)\}} E\left\{\frac{\delta M(\mathbf{X}\mu)}{\delta X}\right\} \tag{4.7}$$

As Firpo et al (2009) exemplifies, the application of UQR is similar to the method used for an OLS regression. To determine the effect at specific quantile φ, the estimation of the RIF for the φth quantile of dependent variable y (log of real hourly earnings) as explained on equation (4.5) and (4.6) respectively. $U_\varphi(p)$ is then estimated using the sample estimate at the unconditional φth quantile. Correspondingly, the density $f_p\{U_\varphi, (p)\}$ at the position $U_\varphi(p)$ is estimated using the uqreg function as shown on equation (4.6) and (4.7). The latter stage is to use OLS to estimate the $\mathrm{RIF}\{y_i, U_\varphi(.), F_p\}$ against the observed covariates, Z.

After the RIF is determined, the UQR can be estimated using a conventional linear regression (LR). In this approach, the corresponding RIF is utilised as the dependent variable y. For instance, in terms of the estimation of interest, earnings (y) is the dependent variable and is a function of the generic skills denoted as (x_i) which is independent variable. The other explanatory variables (z_i) denotes the control variables which includes individual's characteristics such as age and sex etc. Given this, the following model would be estimated using standard OLS:

$$\text{RIF}\{y_i, U_\varphi(.), F_p\} = \alpha_0(\varphi) + \alpha_x(\varphi)x_i + \alpha_z(\varphi)z_i + e_i \tag{4.8}$$

Consistent to the LR model, UQR relies on the assumption that the unobserved component e_I is distributed independently from x_i, and z_i which allows its impact on the conditional expectation of the RIF to be effectively averaged out (Avila, 2020).

Once the independent variables are incorporated in the estimation, the model can be specified as:

$$\text{RIF}\{y_i, U_\varphi(.), F_p\} = \tag{4.9}$$

$$\alpha_0(\varphi) + \alpha_1(\text{literature}) + \alpha_2(\text{numeracy}) + \alpha_3(\text{communication}) + \alpha_4(\text{problem-solving}) + \alpha_5(\text{physical}) + \alpha_6(\text{computer complexity}) + \alpha_z(\varphi)z_i + e_i$$

Where y is the log of real hourly earnings, for individual i. The term $\alpha_0(\varphi)$ corresponds to the unconditional quantile, with φ denoting specific percentile of interest and e_I refers to the disturbance term. The independent variables are denoted as $\alpha_i(.)$ where (.) is the specific generic skill e.g. literature, numeracy etc. In the estimations, the principal focus of interest is the generic skills coefficient. These generic skill coefficients are important and will be subject to the plots for the figures in section 4.5. $\alpha_z(\varphi)z_i$ is a vector of the control variables that have been shown to influence one's hourly earnings. The individuals control various characteristics including age, age square, tenure, occupation, marital status etc.

4.4.3.4. Differences between CQR and UQR

Both CQR and UQR model presents a reliable estimate of a particular target parameter sought by a model. However, it is important to note that while the UQR model's target parameter reflects the impact on outcomes in the entire target population, the CQR model's target parameters pertain to effects on outcomes within specific subgroups defined by the conditioning applied in CQR. Consequently, even if no variables are omitted, the CQR model might exhibit bias when estimating the target parameter of a UQR model.

This method computes the unconditional quantile effect which determines the effect of an explanatory covariate on the outcome of interest at the specific quantile. In this empirical investigation, this implies the implementation of the UQR to gauge the effect of various generic skills on the wage premium across the wage distribution.

Borah and Basu (2013) state there are three main differences between CQR and UQR regressions. Firstly, the estimated impact of a covariate on a particular quantile of outcome remains consistent in both the CQR and UQR regression subject that if no other covariates manipulate the data generated, or if the covariate's effect remains unchanged across various levels of covariate counterparts that affect the data generated.

Secondly, as a covariate's impact on a certain quantile of the outcome changes due to the values of other covariates, the CQR regression provides estimates of the effect on the conditional quantile. This process takes into consideration the average values of all covariates that deviates from the effect on the unconditional quantile.

Thirdly, by incorporating different sets of covariates when there is an involvement of interactive shifts, it causes changes to the estimates that are derived from the CQR model. In contrast, such changes would not impact the estimates obtained from the UQR model, provided that all covariates are considered exogenous.

Therefore, the traditional CQR model estimates the differences between conditional quantile values (Koenker and Bassett Jr, 1978). As Borgen, Haupt and Wiborg (2023) exemplify, when controlling for explanatory variables, the CQR eliminates the impact of these variables and rather, it estimates the independent effects that are located either on conditionally high or low quantiles. As a result, the estimates produced by a CQR model do not provide insights into the position of specific percentile groups located within the overall distribution (Melly and Wüthrich, 2017). Thus, this often restriction limits its utility for estimating quantile effects since the framework often produces results that cannot be applied or interpretable in a policy or population context (Borah and Basu, 2013).

On the contrary, the UQR generates results that are practical for interpretation since it marginalises the effect across the distributions of other covariates in the model. Unlike the CQR, the interpretation of the UQR must be examined from the perspective of a target population to which the regression analyses i.e., the specific percentile position on the wage distribution. Avila and New (2022) state that the UQR coefficients can be examined more intuitively due to the unit's outcome variable quantile is not conditional on the explanatory variables, but rather just the outcome itself (Avila and New, 2022).

The CQR model developed by Koenker and Bassett (1978) cannot address policy issues that depend on the unconditional statistical properties of the response variable. Both quantile regression models are useful in that they allow the explanatory variables to have different impacts across the distribution of the outcome variable. For example, the impact of physical skills on return to wages may be quite different for individuals that are at the 10th percentile of the wage distribution compared with individuals at the 90th percentile. Therefore, the primary point is that when focusing on the group of workers at the 90th percentile, the estimation analyses only the top 10% of the earners in the population, irrespective of their characteristics. However, the CQR model can be used only to assess the impact of a covariate on a quantile, conditional on specific values of the other explanatory variables. For instance, if the CQR focus on the group of individuals at the 90th percentile, the estimation analyses given the characteristics of the individual, they relatively earn the lowest 25%. Therefore, an individual that earns at the top 10% bracket can be in the 25th percentile under the CQR if they earn a relatively low wage for their given characteristic.

Overall, this UQR model yields results which are more detailed which in turn allows for greater interpretability from a policy implication point of view. The benefit of this approach is that the result marginalises the effect over the distributions of other covariates (generic skills) in the model. Consequently, in the analysis, the focus can be on how the wage premium differs at various points of the wage distribution as the importance of a generic skill increases. This approach allows for more detailed policy implications, providing more insightful results by focusing on individuals within specific wage bands.

4.5. Results and discussion

4.5.1. Unconditional quantile regression

The UQR reveals the effect and the magnitude of the effect when the importance of a specific generic skill adjusts for an individual's occupation. This "effect" determines whether the increase in the importance of the generic skill has a positive/negative or no effect on the earnings of an individual at a certain hourly wage level i.e., quantile.

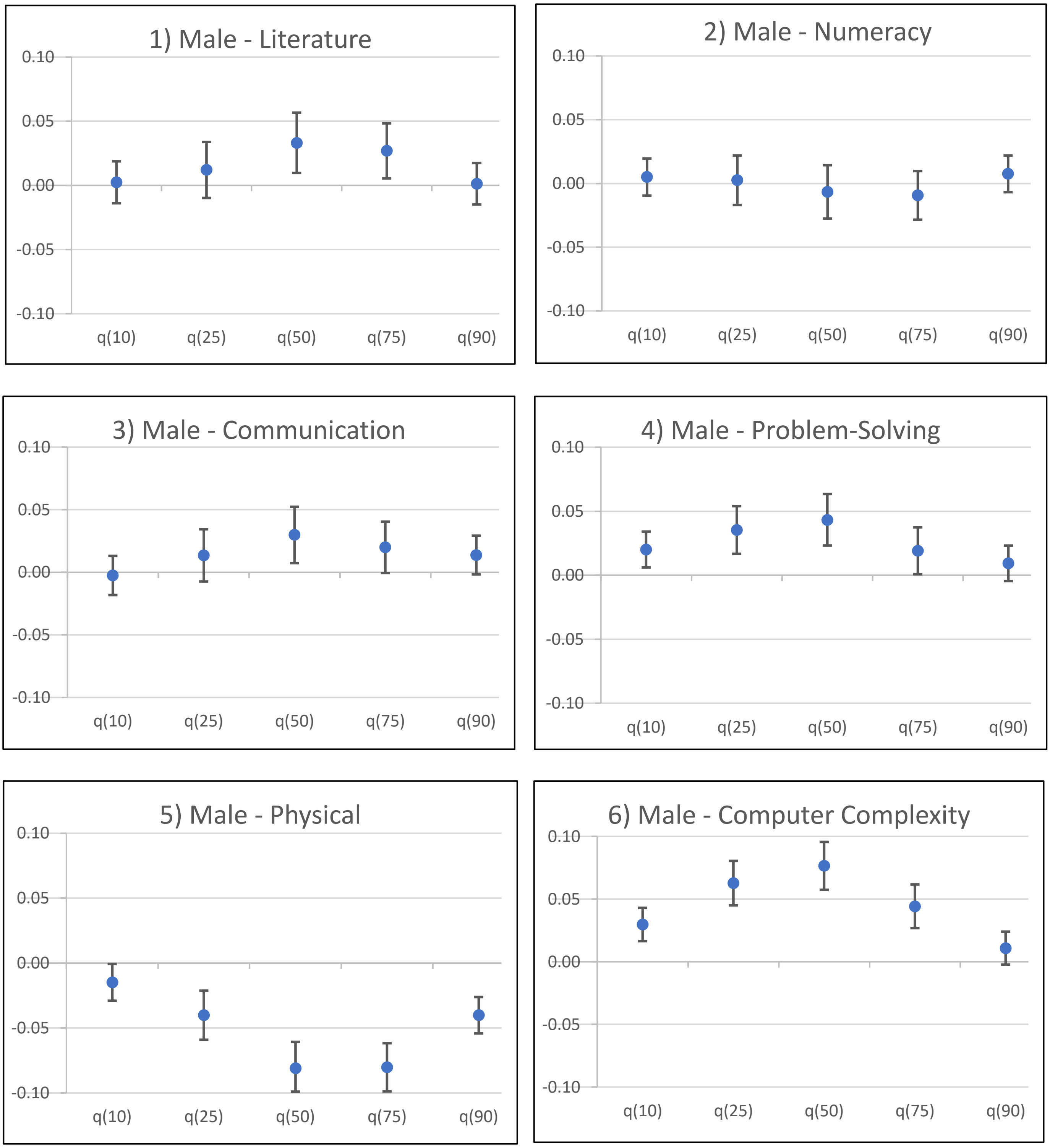
4.5.1.1. Male

Figure 4.3 (4.4) shows the unconditional quantile effects, including 95% confidence intervals, for males (females) across six covariates distinguishing the impacts of various generic skills. The corresponding regression Tables for males (females) are in the appendix Table 4.9 (4.10). In Figure 4.3.1, statistically significant results within the generic skill literature appear only at the 50th and 75th quantiles, with positive odds ratios of 0.033 and 0.027, respectively. However, their 95% confidence interval bands overlap, indicating no statistical differences between the effects at these two quantiles. This suggests that as the importance of literature skill increases in a given occupation, an individual at the 50th quantile does not experience a statistically greater wage premium effect compared to an individual in the 75th quantile range.

Communication skills (Figure 4.3.3) and problem-solving skills (Figure 4.3.4) exhibit a similar pattern to literature skills. These covariates show a consistent, monotonic increase in odds ratio from the lower quantile to the median, reaching a peak, and then gradually decreasing as the quantile rises. However, while communication skills yield three statistically significant results, problem-solving skills encompass four statistically significant findings. Insignificant results with communication skills are positioned at the two lowest quantile points, while problem-solving skills are located at the highest quantile point. The 95% confidence interval bands for communication skills overlap, indicating no statistical differences. Conversely, standard deviation bands at the 10th quantile are notably lower than at the 25th and 50th quantile points. Consequently, there is a statistical increase between the 10th and 25th quantiles, as well as the 10th and 50th quantiles, with p-values of 0.1 and 0.03, respectively.

Figure 4.3

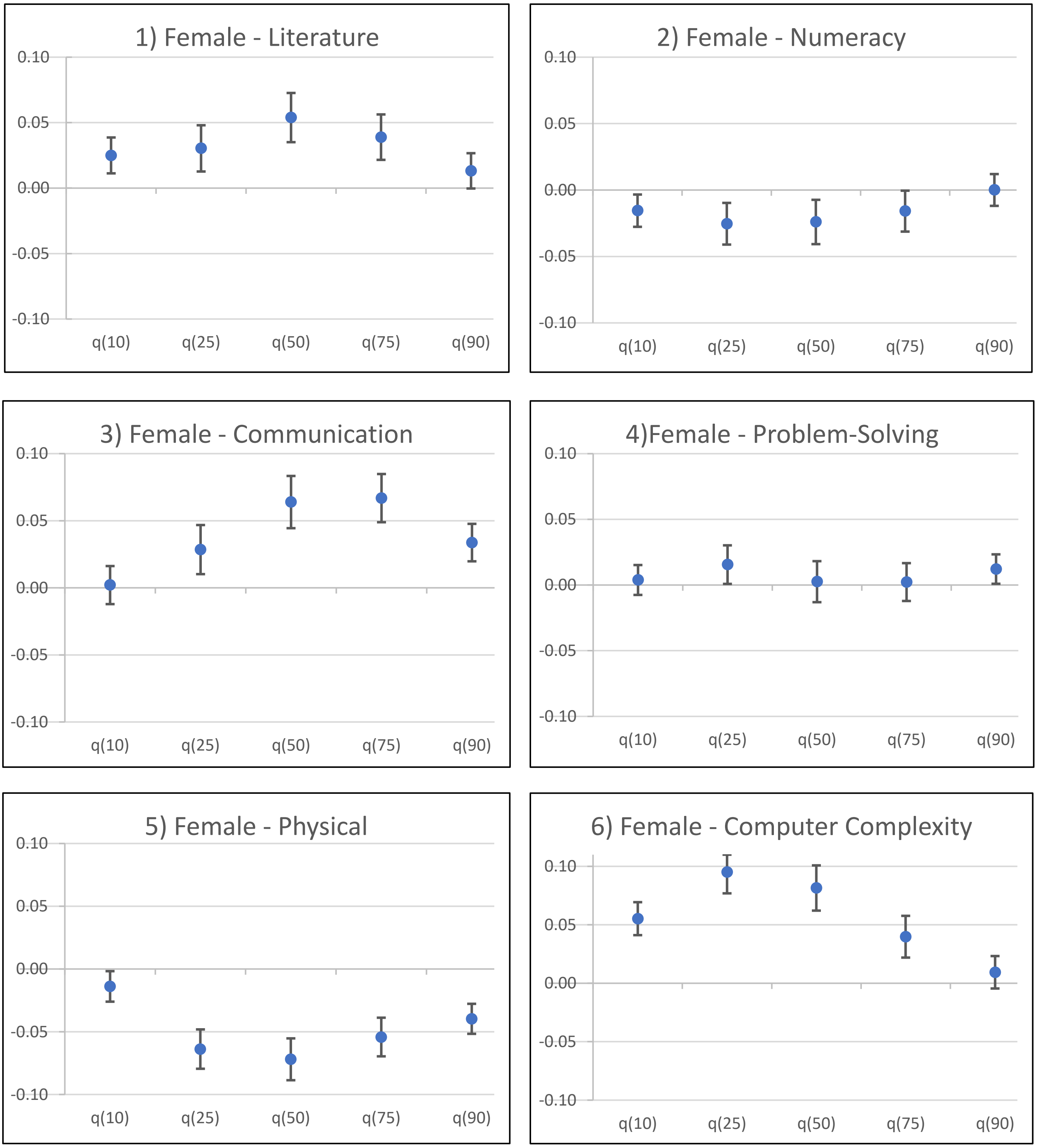
Unconditional quantile regression – Male



Notes: The vertical lines with a bar above and below each of the value points represent the 95% confidence intervals for the coefficient. Figure 4.3 corresponds to Table 4.9 in the Appendix. Table 4.9 shows the returns of the generic skills and complete explanatory variables for males.

Figure 4.4

Unconditional quantile regression – Female



Notes: The vertical lines with a bar above and below each of the value points represent the 95% confidence intervals for the coefficient. Figure 4.4 corresponds to Table 4.10 in the Appendix. Table 4.10 shows the returns of the generic skills and complete explanatory variables for males.

There are similar patterns with the UQRs since the absolute coefficients tend to be closer to zero and with a low level of statistical significance at the two ends of the tails. The coefficient peak is usually located around the median with the highest significance level. The pattern of the UQR is consistent with the remaining generic skills for males. This finding is not surprising as it implies that at the lower and upper quantiles, generic skills do not affect the wage premium significantly. This makes sense as the group of the lowest earners is in occupations whereby the changes in the importance of generic skills are not the main contributor to the hourly wage. In the opposite spectrum, the highest-earning occupations are not likely to significantly increase their wage premium if the importance of an occupation's specific generic skill requirement increases. Individuals at the 90th quantile are often in occupations where their job requires a higher importance of generic skills i.e., several generic skills importance are essential. Therefore, the most profound elasticity for the effect of generic skills upon the wage premium is around the median.

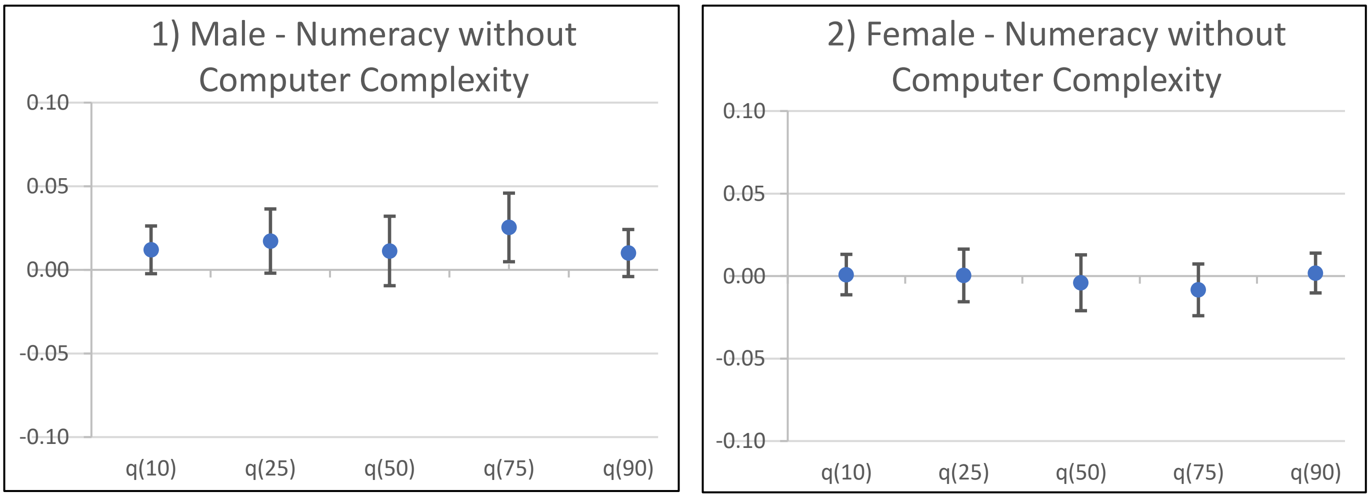
Numeracy skill and physical skill UQR are the only covariates that exhibit a net negative effect as seen in Figures 4.3.2 and 4.3.5. Numeracy skills have two quantile points which experience negative effects, the 50th and 75th quantile points with an odds ratio of -0.007 and -0.009 respectively. Physical skill is an anomaly as all the results are negative at every quantile point, this effect is the most prolific around the median.

The negative effect for physical skill is substantial as at every quantile the odds ratio is negative and are all statistically significant at the 1% level (except for Q10 with a p-value of 0.04). From the lower quantile going towards the median, the effect becomes increasingly negative and troughs at the 50th quantile with a coefficient of -0.81. This odds ratio persists at the 75th quantile and predictably the negative effect significantly reduces at the 90th quantile with a z-value of -3.38 and p-value of 0.00 (from the 75th to 90th quantile). In fact, q(10) is statistically different from all the other quantiles, the same as q(25) except for between q(25) and q(90). Other than between q(50) and q(75), these two quantiles are also statistically different from the counterpart quantiles. This implies conditional on the individual's position on the quantile spectrum, there is a considerable heterogeneous effect between the importance of physical skill in a job and the hourly wage. For instance, between the 10th and 50th quantiles, there is an odds ratio difference of 0.066 points. To conceptualise the difference, the odds ratio is -0.015 at q(10), which implies the negative effect at the median is over five times greater than at q(50).

Consistent with the paper by Francis Green (2003), they also found that physical skill negatively affects the wage premium. The focal reason for this effect is that manual skills tend to be negatively correlated with other observed and unobserved skills, those of which are generally positively valued skills. Evidence of this can be observed in Table 4.3 since the Pearson correlation coefficient for physical skills is generally negatively correlated with the other generic skills. In turn, individuals who work in occupations where the use of physical skills is important, workers usually do not use other generic skills which are regarded as more highly valued skills. Another explanation is that the labour market considers manual skills such as physical stamina to receive a relatively low or even zero supply price.

Figure 4.5

Unconditional quantile regression - Numeracy without computer complexity



Notes: The variable computer complexity is excluded when estimating the returns for numeracy skills. The purpose is to observe whether computer complexity significantly affects the returns to numeracy skills.

Curiosity may arise given that numerical skills incur a small negative premium, especially since evidence from previous papers suggests the increasing importance of mathematical skills (Dolton and Vignoles, 1999). However, as Figure 4.3.2 shows, all the quantile points are statistically insignificant which implies that the importance of numerical skills does not significantly affect an occupation's hourly rate. Francis Green (2003) explains this enigma arises since numerical skills and computing skills are highly correlated. Table 4.3 illustrates that numerical skills have the highest Pearson correlation with computer complexity and vice versa. Computer complexity exhibits the greatest positive wage premium in comparison to its generic skill counterparts. Therefore, due to the strong correlation between computer complexity and numerical skills, this dramatically reduces by mitigating the positive effect that can be observed econometrically between numerical skills and the wage premium.

Robustness checks on the numeracy effect can be configured by investigating the wage premium when computer complexity is omitted in the UQR. The assumption would be that numeracy skill would generate a more positive effect in comparison to the numeracy skill effect from the non-augmented regression since computer complexity skill would not mitigate the positive effect. As predicted, the augmented UQR shown in Figure 4.5.1 illustrates that the numeracy skill possesses a more positive effect when the covariate computer complexity is omitted. All the coefficients are positive and three of the quantile points record a statistically significant result. This confirms that in the absence of computer complexity in the UQR, numerical skill yields a positive effect on the wage premium. However, it is peculiar of the pronounced decline in the odds ratio and the reduced significance at the median.

Finally, to examine the effect of computers on wage premia, a computer complexity variable is incorporated in the UQR. This variable includes the levels of computing complexity with the following indicators "do not use at all, straightforward, moderate, complex and advanced" where the base category is those who do not use computers at all. As shown in Figure 4.3.6, a higher level of computing complexity in an occupation induces a higher wage premium, in fact, compared to its counterparts, computer complexity possesses a considerably greater positive wage premium. For instance, at every quantile point, computer complexity consists of the highest positive coefficient compared to every covariate, and except for q(90), all the p-values are significant at the 1% statistical level. Interestingly, similar to problem-solving, the effect is stronger at the lower quantiles than at the upper quantile. This implies that there is a higher propensity for an increased wage premium at the lower-earning spectrum compared to high earners. There are many factors for this to occur, one of which could be that it is common knowledge that as a skill increases, the wage increases at a slower rate due to the law of diminishing returns. This is true for high-earning individuals, thus, for further perpetual growth in earnings to occur, an individual would potentially require developing other skills such as interpersonal and/or soft skills, etc.

The paper by Green (2003) suggests that utilising a "computer usage" dummy to capture wage effects is likely to be misleading since the indicator is too simple. The findings confirm this as two individuals who report computer usage as essential in a workplace do not necessarily implicate similar earnings. For example, an individual who operates a computer for programming purposes would likely receive a considerably greater wage premium in comparison to an individual using a computer to answer calls in a call centre. Distinguishing low-skilled jobs is important for policymakers in an attempt to alleviate poverty.

Johnson (2007) found that job prospects progression differs as some occupations allow for wage increases over the evolution of employment, while in other jobs wage propensity is limited. Johnson analysed the employment survey between 1997 and 2004 and discovered that workers with greater relative skills such as computer complexity are more inclined to jobs with greater skill requirements and the ability to attain greater wage growth. As witnessed, this is consistent with the findings as the propensity of the change in the wage premium is very elastic for computer complexity. For example, the odds ratio at q(10) is statistically different from all the other quantiles which suggests that the other quantiles affect the wage premium significantly more. In fact, the median has the highest effect on the wage premium with an odds ratio of 0.076. Moreover, it is evident that q(50) and q(75) are statistically different as the standard deviation band does not overlap each other as seen in Figure 4.5, this is reliable with a z-value of 2.45 and a p-value of 0.01. This is not the case between q(50) and q(25) as the lower standard deviation band for q(50) intersects with the upper standard deviation of q(25).

4.5.1.2. Female

Figure 4.4.1 to 4.4.6 presents the UQR for females. The purpose of this investigation is to observe whether generic skills affect the wage premium differently between genders. Overall, the returns to the various generic skills between the occupations for males and females are relatively similar with comparable trends. However, there are several key differences between the genders, though it should be noted that there is a considerable unexplained covariate for the difference in earnings between genders since not all factors can be accounted for due to the complexity of the wage composition.

For example, there remains gender difference in communication and problem-solving skills since they both demonstrate significant deviation in wage premium compared to the other generic skills. For instance, men experience a relatively flat wage premium effect with communication. In contrast, women enjoy a high and significant premium, especially in the middle and upper part of the wage distribution. Observing Figure 4.4.3, the odds ratio at q(50) and q(75) is 0.064 and 0.067, respectively at a 1% significance level. This indicates that the female wage premium effect from communication skills is over twice as much compared to males.

According to Beaudry and Lewis (2014), females possess an inherent advantage in leadership positions within public relations due to their inclination towards relationship-oriented traits. These situational leadership advantages are further impelled by the gender differences in communication styles and influence tactics. This explains the higher wage premium effect for females in terms of communication skills as female leaders exhibit a greater degree of intimate and relational engagement during conversations. Therefore, females are particularly well-matched for occupations in human resources or public relations since occupations, where communication is important, require a core duty that involves effective communication, attentive listening, and addressing the diverse needs of individuals.

Furthermore, the quantiles q(10), q(50) and q(75) are not significantly different to zero. Also, the two quantiles which are positive and significant, q(25) and q(90) are not statistically different to the other three quantiles as Figure 4.4.4 shows. This suggests that the wage premium effect for females derived from problem-solving skills is relatively flat.

Comparing the outcome by males shown in Figure 4.3.4, there is a considerable discrepancy between the wage premium effect which could be caused by gender differences. The greatest difference between the genders is at the lower and the middle part of the wage distribution. This implies that as the importance of problem-solving heightens, men receive a greater wage premium effect compared to women, especially at the mean of the wage distribution since the odds ratio is 0.043 for males and 0.003 for females.

Problem-solving requires an individual the ability to identify problems, brainstorm, analyse and implement the best solutions. Therefore, to be effective at problem-solving, one must be assertive and proactive in seeking solutions. Heckman et al (2010) expressed that highly agreeable individuals are frequently characterised as compassionate, polite, and kind which may appear beneficial to one's career success. However, Heckman et al found that highly agreeable women, receive a wage penalty. In their research, the men's agreeability was insignificant which resulted in a significant wage advantage. This is consistent with the finding that men experience a greater wage premium when problem-solving is of interest due to the natural tendency to be more decisive and firmer.

The remaining UQR for females possesses a similar trend to males. For instance, both Figure 4.3.1 and 4.4.1 coefficient increases from the lower quantile to the median and monotonically tapers off. Although, it is noticeable that the returns to literature are considerably greater for females than males. This is true at every quantile and unlike the male counterparts, the female odds ratio is all significant at the 1% statistical level (except at the 90th quantile). According to Barrett and Staneva (2017), different traits and cognitive skills can help explain the gender pay gap by finding significant discrepancies in how men and women are rewarded or penalised. In line with the outcomes, Barrett and Staneva found that women see a larger benefit for their word reading ability skills compared to men. Their studies found despite no significant differences in cognitive skill levels between the genders, females are rewarded more than males. For instance, in the national word reading test which evaluates premorbid intelligence, females at the upper quantiles receive a significantly higher reward of 8% wage premium compared to males only receiving 5%. Therefore, this reinforces that males and females both benefit from a wage premium as the importance of literature skills increases, although the effect of the increased wage premium is greater for females.

The female UQR for numeracy skills illustrates a similar trend to their male counterpart. All the coefficients in Figure 4.4.2 are negative and statistically significant (except at the 90th quantile) whereas none of the coefficients for males are statistically significant as shown in Figure 4.3.2. This implies that numeracy skills for females have a negative impact on returns, while the result for males is insignificant. However, in the augmented regression where computer complexity is omitted (Figure 4.5.2), female numeracy skill induces no effect on the wage premium as all the coefficients are statistically insignificant. This trend is somewhat similar to the UQR for problem-solving. Both Figure 4.3.4 and 4.4.4 exhibits a relatively flat trend situated around the zero origin. Therefore, there is a reduced effect of problem-solving on the wage premium for females than males with only two coefficients providing significant results.

The paper by Christl and Köppl-Turyna (2020) analysed the Austrian gender wage gap and the role of skills by utilising the PIAAC and also found there are substantial differences in the skill use between genders. This is especially the case for numeric skills as males tend to use this skill on average more often than females. In line with the results obtained, Christl and Köppl-Turyna found that numerical skills show a wage premium for males, while for females the returns are around zero.

Interestingly, Barrett and Staneva (2017) suggest that the changes in the wage premium with respect to the importance of a generic skill in an occupation could lead to men and women making different career choices. This suggests that the relationship between gender, specific generic skills, and performance is not straightforward. It is essential to recognise that men and women make different career choices, which significantly impacts income and the wage premium.

The effects of physical skills seem to be rather homogenous to the wage premium irrespective of gender. Both genders experience a negative wage premium at every quantile. However, the negative effect on returns for females is considerably less at the 50th and 75th quantile in comparison to males. On the other hand, the coefficient at the 25th quantile is significantly more negative for females than males with a z-value of -1.89 and a p-value of 0.03. Overall, the negative wage effect of physical skills is less compelling for females around the middle part of the distribution than for males. At the lowest and highest part of the distribution, both gender returns by physical skills are virtually the same.

Furthermore, the UQR for females in terms of computer complexity comprises a higher wage premium effect than males despite previous research concluding that males attribute to STEM subjects more often than females. Beaudry and Lewis (2014), argue that in the past decades, computer technology has raised the value of cognitive and interpersonal skills over physical skills, thus, on average these benefits are more so for females. Economists such as Finis Welch note that the male-female wage gap may be related to more general changes in how the market values people with different skills. As such, Bacolod and Blum noticed that women are more likely to use computers at work than men, as well as to work in occupations and industries that have greater benefit from computerisation. This is especially the case at the lower level of the earning distribution since the coefficient at the 25th quantile is greater than at the 50th quantile as illustrated in Figure 4.4.6 (although the figures are not significantly different with a z-value of 1.01 and p-value of 0.16). However, at the upper end of the earning distribution, both genders have a similar wage premium effect and can even be argued that men slightly perform better.

4.5.1.3. UQR: White British

The examination has focused on how each generic skill influences the wage premium and whether there is any disparity in results between genders. It is intriguing to observe the influence of ethnicity on generic skills and, if present, how it impacts the wage premium across different quantiles. As such, Figure 4.6.1 – 4.6.6 (4.7.1 – 4.7.6) illustrates the UQR of the wage premium throughout the distribution for each generic skill subject to solely including White (Non-White) observations. In comparison to Figure 4.3.1 and Figure 4.4.1, the UQR of Figure 4.6.1 produces relatively similar results. This is predicted since the majority of the observations captured by Figures 4.3.1 and 4.4.1 are represented by the White ethnicity, thus it is expected that the results would be similar and all the result except at q(90) is statistically significant at the 1% level. However, the literature generic skill of the White ethnic group shows that the wage premium is flatter across the earning distribution compared to both genders since they experience a greater peak around the median.

Figure 4.6

UQR – White British

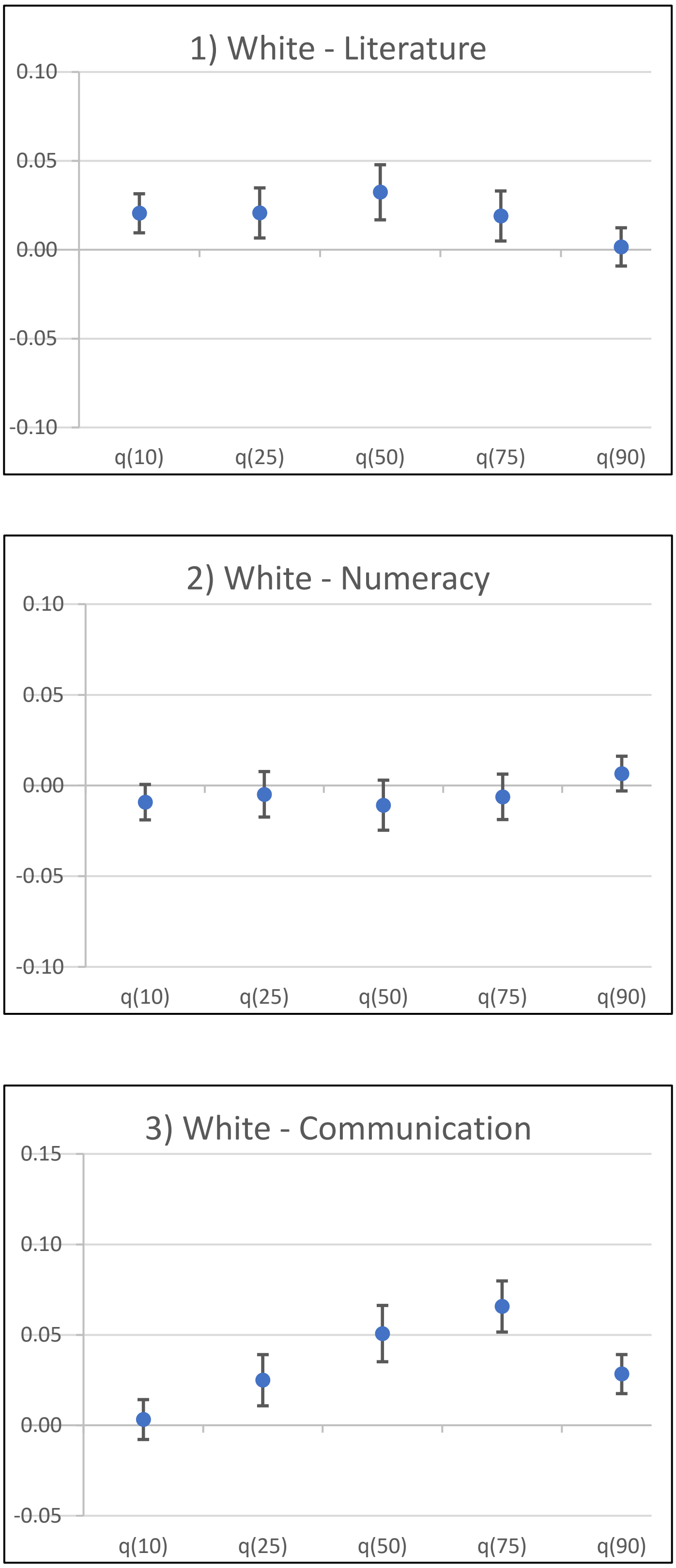
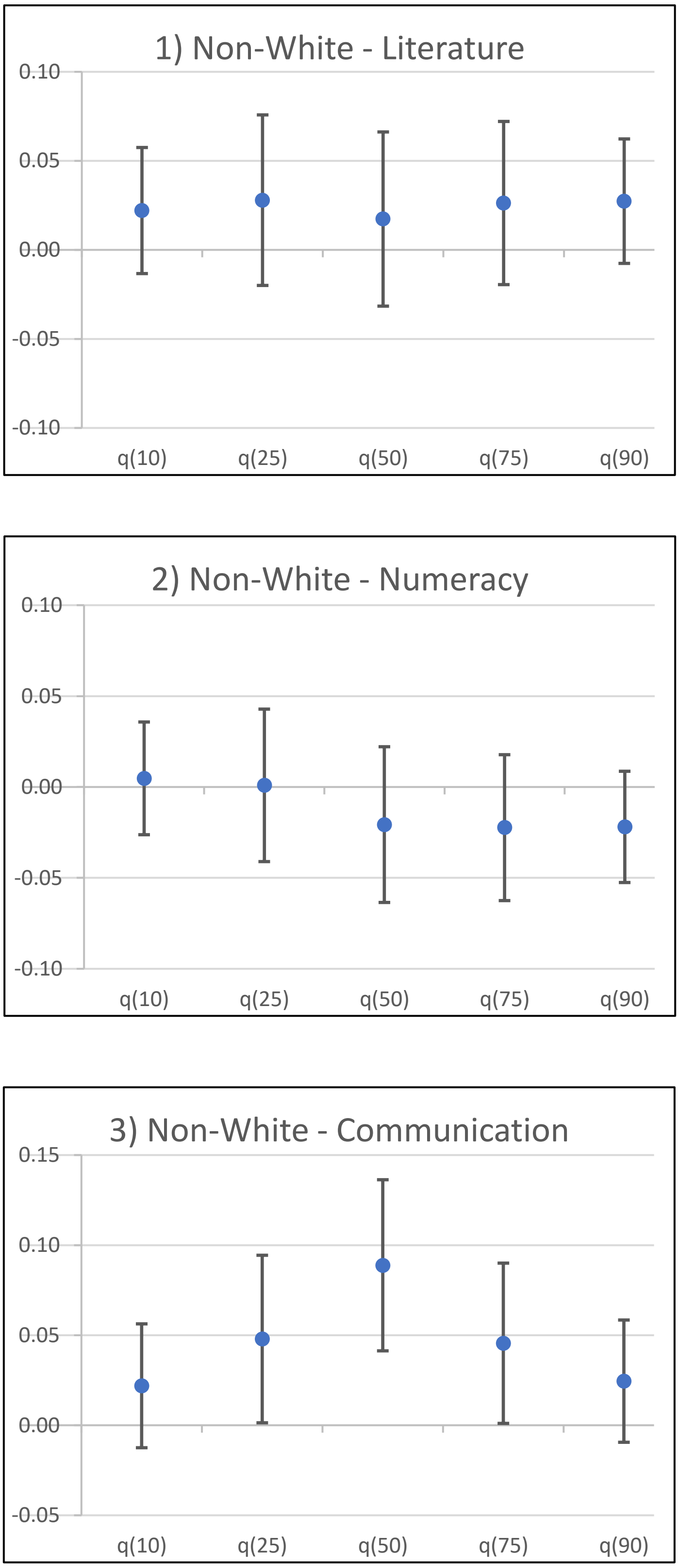
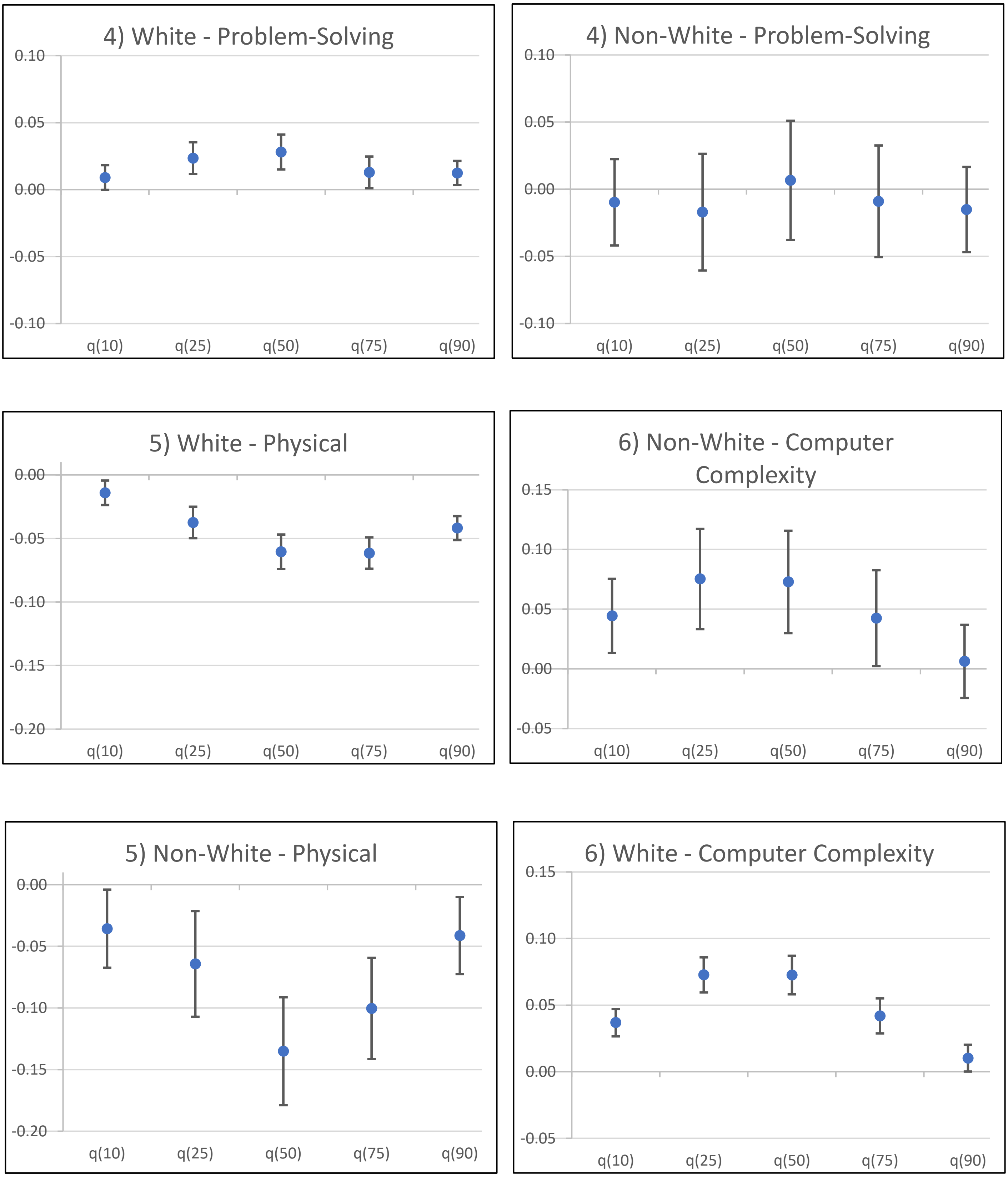


Figure 4.7

UQR – Non-White





Notes: Figures 4.6 and 4.7 correspond to Table 4.11 and Table 4.12 respectively in the Appendix. Table 4.11 and 4.12 shows the returns of the generic skills and complete explanatory variables for White and Non-White ethnic group respectively.

The reasoning of this is because as previously explored, females' wage premium is in general greater than males when focusing on literature skills. Therefore, collectively the results offset each other which in turn the results become more constant throughout the wage distribution.

The contrasting generic skills show the wage premium results for the White ethnic group are consistent with the literature skill in the sense that most of the results are similar to the previous gender results that correspond to their respective generic skill. For instance, in the case of the wage premium for communication skills (Figure 4.6.3), the coefficients tend to be between the values of the two gender results i.e., Figure 4.3.3 and Figure 4.4.3. This is the case for the wage premiums for the other generic skills. It should be noted that for numeracy skills, alike the male results, the coefficients are not statistically significant. With respect to the other generic skills, the p-values are all significant except for communication at the 10th quantile. With the exception of numeracy skills, the White ethnic groups generally produce more statistically significant results than both genders. This is driven by excluding the ethnic minorities' results as by including the observations in the UQR, the results if offset and become distorted i.e., reduce significantly.

4.5.1.4. UQR: Non – White

There is greater interest in investigating the ethnic minorities' wage premium as the results are less predictable. Given the relatively limited observations for ethnic minorities, the approach involves grouping these minority groups and categorising them collectively as the Non-White group. In terms of literature skills, the odds ratio for the Non-White group are all insignificant, thus there is evidence that the coefficients are not different to zero. However, Figure 4.7.1 shows that the wage premium for the Non-White group is greater than the White group. This is true across the entire wage distribution except at the 50th quantile. This suggests that in general, the ethnic minorities earn more than the White ethnic group as the importance of literature skills increases. Since Figure 4.7.1 shows a consistently higher odds ratio than Figure 4.6.1, the result correlation is not necessarily meaningless.

On contrary to the White group, the ethnic minorities wage premium with respect to the numeracy skills shows a negative correlation. As shown in Figure 4.7.2, when moving up the wage distribution, the wage premium decreases. However, the magnitude of the coefficients is not large and are statistically insignificant which is similar to the White group. This occurrence is consistent with the problem-solving skill of the Non-White group as all the coefficients are not significant at any point of the wage distribution. Also, Figure 4.7.4 shows that the coefficients are relatively close to zero as well as the upper and lower standard deviation limits are all positive and negative respectively. Therefore, the result suggests that the importance of numeracy skills and problem-solving skills for the Non-White group, in general, does not have a significant effect on the wage premium of an individual.

The wage premium related to communication skills as illustrated in Figure 4.7.3 exhibits a pattern akin to that of males in Figure 4.3.3, with the coefficient peaking at the median and the troughs at both the lower and upper extremes of the wage distribution. As such, q(10) and q(90) coefficients are not statistically significant. The wage premium at q(25) and q(50) are higher than the White ethnic group, although, at the higher wage distribution, the White ethnic group outperforms the Non-White ethnic group, especially at q(75) with a greater odds ratio of 0.011. This suggests that at the lower and median of the wage distribution, the Non-White group tend to receive a higher wage premium and at the higher wage distribution, the White ethnic group performs better. This implies that among individuals with lower wages, having stronger communication skills results in a higher wage increase for the Non-White ethnic group compared to the White ethnic group.

However, Okafor et al. (2023) found that 93% of employers believe interpersonal and communication skills are "very important" or "essential". According to Okafor et al., individuals with strong interpersonal skills are 14% more likely to earn a top-quintile income. This translates at the upper quantile especially at q(75) as the White ethnic group on average is expected to communicate more proficiently compared to the ethnic minority counterparts. This effect tapers at q(90) since the odds ratio is similar for both groups. This could suggest those who earn the highest wages typically possess strong communication skills, resulting in a similar impact on wage changes between the White and ethnic minority groups when the importance of communication skills increases.

Figure 4.7.5 illustrates the wage premium for the Non-White ethnic group subject to changes in the importance of physical skills. It is evident that consistent with the previous graphs for physical skills, all the coefficients are negative at every quantile across the wage distribution. In fact, the wage premium for the Non-White ethnic group is lower than the previous groups (Male, Female and White groups) at every quantile point. All the findings are statistically significant, indicating compelling evidence that the Non-White group exhibits the poorest performance. The wage penalty for physical skills is most pronounced at the median, and it is statistically inferior to all other quantiles, except for q(75), where the z-value is 1.14, and the p-value is 0.1 Given that individuals at q(75) experience a notable wage penalty, they are statistically in a worse position than those at q(10) and q(90).

A potential reason for the greater wage penalty compared to its counterparts is the type of job taken within an occupation. For instance, the Non-White ethnic group which requires a relatively higher emphasis on physical skills may often be employed in physically demanding, labour-intensive occupations. which require heavy labour-intensive jobs. Conversely, the White ethnic group, on average, may be engaged in professions that require a higher level of proficiency in various physical skills which entails higher wages such as operating advanced machinery and technology.

The wage premium coefficients between the Non-White and White ethnic groups for computer complexity are relatively similar. Similar to the other three computer complexity findings, except for the highest point of the wage distribution, all the coefficients are statistically significant. The wage premium coefficients for computer complexity are comparable between the Non-White and White ethnic groups. Much like the results for the other three findings of computer complexity, all the coefficients are statistically significant, except for the top end of the wage distribution.

4.5.2. UQR: Considerations of the eigenvectors during the PCA

The PCA employed to calculate the UQR outcomes depicted in Figures 4.6 and 4.7 utilises the dominant eigenvalues. As detailed in the methodology (Section 4.3.5.1), the selection of the first principal component is based on its possession of the largest eigenvalue, thereby explaining the most significant variance. With the exception of communication skills, each generic skill experiences only one component with an eigenvalue exceeding one. Hence, for generic skills with a singular eigenvalue greater than one, the highest principal component is selected during estimation.

However, communication skills have four components which experience an eigenvalue greater than one. This suggests that there are several principal components that capture significant variance of the generic skill.[19] There is argument of whether to retain these components as it can reduce the dimensionality of the dataset while preserving most of the information. The alternative components greater than one can provide a more comprehensive understanding of the data distribution. Therefore, as a robustness check, further estimations will be conducted on the returns associated with generic skills. This will involve using the three eigenvalues that exceeds one which is generated for the PCA of communication skills. Further information provided in the appendix, Figure 4.7.

4.5.3. UQR: Graduates

It is presumed that the influence of generic skills on wage returns for graduates is less significant compared to non-graduate peers. Table 4.4 exclusively examines a subset of graduates, showcasing a UQR of the impact of generic skills on wage premiums. Numerous coefficients are statistically significant which indicates there is sufficient data for analysis. However, upon inclusion of ethnic groups, all findings become insignificant due to insufficient sample size, thereby reducing statistical power. Therefore, for this estimate, the ethnic minority variable is excluded.

Table 4.5. presents a UQR estimate of the wage equation for graduates while accounting for the subject chosen during the individuals university studies. In comparison to Table 4.4., literacy skills contribute to an increase in the wage premium. However, Table 4.6 indicates that this change is statistically significant only at the 25th quantile, with a z-value of 2.67 and a p-value of 0.00. Apart from literacy skills, the impact of other generic skills either remains relatively consistent or results in a decrease in the wage premium. For example, the inclusion of graduates' subject choices in the estimation decreases the wage premium for numerical skills. However, there are no significant wage changes when compared to the results where graduates' subject choices are excluded, as shown in Table 4.6.

[19] The first principal component has an eigenvalue of 4.59 which is significantly greater than the other three principal component with an eigenvalue of 1.20, 1.12 and 1.02 respectively from largest to smallest. As a result, the first principal component is chosen as it significantly explains the most variance of communication skills. Therefore, the component retains the largest amount of information possible from the original data.

Table 4.4

Regression of graduate wages on generic skills and controls (Excluding university subject choices)

Generic Skills	Q10	Q25	Q50	Q75	Q90
Literacy Skills	0.071***	0.014	0.012	-0.009	0.006
	(0.020)	(0.027)	(0.033)	(0.029)	(0.021)
Numerical Skills	-0.042**	-0.037	0.020***	0.019	0.016
	(0.020)	(0.026)	(0.032)	(0.029)	(0.021)
Communication Skills	0.032*	0.065***	0.083***	0.069**	0.050**
	(0.020)	(0.026)	(0.033)	(0.030)	(0.021)
Problem-Solving Skills	0.030	0.050*	0.055	0.093***	0.041**
	(0.020)	(0.026)	(0.033)	(0.030)	(0.021)
Physical Skills	-0.009	-0.014	-0.071**	-0.047*	-0.043**
	(0.021)	(0.027)	(0.034)	(0.030)	(0.022)
Computer Complexity	0.073***	0.107***	0.043	0.017	-0.018
	(0.021)	(0.028)	(0.035)	(0.031)	(0.023)
Constant	**0.621***	**-0.277**	**-1.367****	**-0.971**	**-0.777**
	(0.337)	**(0.443)**	**(0.557)**	**(0.500)**	**(0.361)**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. Q10 – Q90 denotes the quantile of interest. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places.

Furthermore, there is a noticeable decrease in the wage premium for communication skills in Table 4.5. This reduction is substantial across all percentile points, although the only statistically significant result is observed at the 25th quantile. The p-value of the z-test indicates significant changes to the wage premium at every quantile as demonstrated in Table 4.6. This is consistent with Chawner and Oliver's (2013) findings, suggesting that many universities and students prioritise technical contents of their degree subjects as opposed of the improvement of their generic skills. As a result, the inclusion of the degree subject is expected to diminish the impact of the wage premium associated with generic skills which is evident in the case of communication skills.

Table 4.5

Regression of graduate wages on generic skills and controls (Including university subject choices)

Generic Skills	**Q10**	**Q25**	**Q50**	**Q75**	**Q90**
Literacy Skills	0.071*** (0.020)	0.014 (0.026)	0.012 (0.033)	-0.009 (0.030)	0.007 (0.021)
Numerical Skills	-0.042*** (0.020)	-0.037* (0.026)	0.201 (0.032)	0.019 (0.029)	0.016 (0.021)
Communication Skills	0.032 (0.020)	0.065** (0.026)	0.084* (0.033)	0.069** (0.030)	0.050** (0.021)
Problem-Solving Skills	0.030 (0.020)	0.050* (0.026)	0.055* (0.033)	0.093*** (0.030)	0.041* (0.021)
Physical Skills	-0.009 (0.020)	-0.014 (0.027)	-0.071** (0.040)	-0.047 (0.030)	-0.043** (0.022)
Computer Complexity	0.073*** (0.021)	0.107*** (0.028)	0.043 (0.035)	0.017 (0.031)	-0.018 (0.023)
Constant	**0.621* (0.338)**	**-0.277 (0.443)**	**-1.367* (0.557)**	**-0.971 (0.500)**	**-0.777** (0.361)**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. Q10 – Q90 denotes the quantile of interest. The dependent variable is real earnings in 2010 prices. All entries outside (inside) parentheses rounded to three (three) decimal places.

For the three remaining generic skills (problem-solving, physical, and computer complexity), the coefficients remain largely consistent between Table 4.4. and Table 4.5. Conducting z-tests for these three generic skills reveals no significant changes in the wage premium. In summary, in conjunction with the findings regarding communication skills, this suggests that the influence of generic skills on the wage premium diminishes for university graduates.

Table 4.6

P-values of Z-test between estimates of graduates with and without including university subject choices

Generic Skills	**Q10**	**Q25**	**Q50**	**Q75**	**Q90**
Literacy Skills	0.50	0.00	0.11	0.39	0.40
Numerical Skills	0.20	0.38	0.35	0.29	0.28
Communication Skills	0.07	0.00	0.00	0.02	0.05
Problem-Solving Skills	0.36	0.50	0.45	0.41	0.48
Physical Skills	0.23	0.44	0.43	0.37	0.41
Computer Complexity	0.25	0.47	0.49	0.36	0.29

Note: Table 4.6 only shows the p-values of the z-test. Only communication skills indicate statistically significant results. Q10 – Q90 denotes the quantile of interest.

4.6. Conclusion

4.6.1. Summary

In summary, this study reveals key insights into the impact of generic skills on wage premiums within the UK labour market, highlighting both gender and ethnic disparities. The

Unconditional Quantile Regression analysis shows that, while males and females follow similar wage patterns across skill categories, females tend to experience a more positive overall wage premium from generic skills than males. Men outperform women in specific skills, notably numeracy and problem-solving, which may reflect occupational preferences and the structural traits of certain job sectors. Interestingly, the wage premium gap between men and women appears to have narrowed compared to earlier studies, like Green et al. (2012), likely due to recent policy interventions aimed at reducing gender inequality. However, men still enjoy an advantage in fields that emphasize analytical rigor and problem-solving, which may be linked to occupational traits and gendered preferences for competitive roles. Research by Cortes and Pan (2018) suggests that men are more likely to select occupations where these traits are valued, which could contribute to persistent wage differences across gendered lines.

Conversely, women show higher wage premiums in literacy and communication skills, which may be linked to leadership styles often associated with femininity, such as people-oriented management approaches. These findings support theories, such as those posited by Catalyst (2007), suggesting that the effectiveness of leadership varies based on occupational roles, with feminine leadership styles benefiting women in fields like human resources and public relations. Meanwhile, roles that emphasize goal-oriented, masculine leadership tend to favour men, especially in sectors like finance, technology, and R&D. Additionally, a broader analysis of cognitive and psychological traits indicates a predictive relationship between personality attributes and occupational choices, highlighting that gender-based preferences may contribute to differences in career paths, earnings, and measures of success within various industries.

This study also highlights the significant impact of computer skills on wage premiums, which stands out as the most valuable generic skill across both genders and all wage quantiles. Advanced and complex computer usage yields a substantially higher wage premium than basic computer tasks, underscoring the growing demand for digital competencies in the modern workplace. Conversely, physical skills continue to attract a negative wage premium, reflecting their diminishing value in an increasingly technology-driven labour market.

Moreover, the wage gap between ethnic minorities and their white counterparts is a notable area of concern, particularly within skill areas like problem-solving and physical tasks. This gap is partially influenced by occupational segregation, where white British workers are more likely to occupy skilled roles involving heavy machinery, while ethnic minorities are often employed in repetitive, labour-intensive roles with lower skill requirements. According to Cramer (2022), male employees from ethnic minority backgrounds in the UK earn approximately 11% less than equivalent white male employees, with a 7% disparity observed among female employees. Despite efforts to close this ethnic wage gap, these findings suggest that it persists, especially in skills and occupations that are undervalued or less compensated.

4.6.2. Policy implications

The findings carry substantial policy implications. Government bodies and educational institutions might consider prioritising the integration of these generic skills into both academic curricula and vocational training programs. By focusing on skills that promote adaptability and critical thinking, the UK can better prepare workers to succeed in a fast-evolving economy where technological change is constant. Policies aimed at skill development could also help bridge the gaps that traditional education metrics fail to address, such as disparities in opportunities for certain groups who may face barriers in achieving conventional qualifications.

Employers should foster these skills through workplace training and development programs. Investing in skill-building initiatives not only improves individual outcomes but can also lead to broader economic benefits, as a highly skilled workforce is likely to drive productivity and innovation. This approach aligns with ongoing efforts to solve the UK's "Productivity Puzzle," as equipping workers with the skills to perform effectively and creatively can contribute to more robust economic growth.

Furthermore, enhancing access to skill development opportunities can be a tool for reducing wage inequalities across gender and ethnic groups. The study found disparities in the returns on certain skills between different demographic groups, underscoring the need for policies that address these inequities. Programs targeting these gaps in skill acquisition and returns could help make the labour market more inclusive, giving underrepresented groups a fairer chance at higher-paying roles.

Overall, the findings suggest that a coordinated policy approach to skill development focusing on generic, adaptable skills can create a more resilient, equitable, and competitive labour market. The UK's economy stands to benefit from policies that support both individual empowerment through skill acquisition and collective progress toward a more inclusive, skilled workforce.

4.6.3. Limitations

Firstly, the analysis relies on quantile regression techniques to explore the relationship between generic skills and wage premiums. While quantile regression offers advantages over traditional mean-based regression, such as the ability to capture heterogeneous effects across different segments of the wage distribution, it also has limitations. For instance, quantile regression assumes linearity between the covariates and the outcome variable, which may not always hold true in practice. Nonlinear relationships between generic skills and wage premiums could lead to biased estimates and undermine the validity of the findings.

Secondly, the study examines only a select set of generic skills, such as literature, communication, problem-solving, numeracy, physical skills, and computer complexity. While these skills are undoubtedly important, they represent only a subset of the broader skill set that individuals bring to the labour market. Omitting other relevant skills, such as interpersonal skills, adaptability, and creativity, could limit the comprehensiveness of the analysis and potentially overlook important determinants of wage differentials.

Additionally, the analysis focuses solely on the direct relationship between generic skills and wage premiums, without considering potential mediating or moderating factors. For example, the study does not account for the role of job characteristics, industry differences, or firm-specific factors that could influence the observed associations. Ignoring these contextual factors may lead to an incomplete understanding of the mechanisms underlying wage differentials and limit the generalizability of the findings to real-world settings.

Furthermore, the study relies on cross-sectional data, which provides a snapshot of the relationship between generic skills and wage premiums at a single point in time. Cross-sectional data may suffer from issues such as omitted variable bias, reverse causality, and unobserved heterogeneity, which could confound the estimated effects and undermine the robustness of the results. Longitudinal data that track individuals' career trajectories over time would offer more insights into the dynamic nature of skill-wage relationships and help establish causal relationships.

Finally, the analysis focuses exclusively on the UK context, which may limit the generalisability of the findings to other countries or regions with different labour market structures, institutional arrangements, and socio-economic dynamics. Cultural differences, institutional factors, and policy interventions could shape the relationship between generic skills and wage premiums differently in other settings, highlighting the need for caution when extrapolating the results beyond the UK context.

4.6.4. Future Research

One potential area for further investigation is the intersectionality of gender and generic skills on wage premiums. While the study already explores differences between males and females, deeper analysis could delve into how factors like race, ethnicity, or educational background intersect with gender to shape wage differentials[20]. Understanding these intersections can help policymakers develop more targeted interventions to address disparities in earnings and opportunities.

Moreover, the study hints at the importance of specific industries or occupations in shaping wage premiums for different skills. Future research could focus on conducting sector-specific analyses to uncover how the demand for generic skills varies across industries and occupations. This could provide valuable insights for individuals seeking to maximize their earning potential through skill development and career choices.

---

[20] In this research, the limiting factor to examine race and ethnicity is the lack of sound sample size, especially for the ethnic minority groups.

Another avenue for research is exploring the long-term effects of skill development initiatives on wage premiums. By following individuals over time and evaluating the impact of interventions such as training programs or educational reforms, researchers can assess how investments in skill development translate into improved earning potential and socio-economic mobility.

Additionally, there is scope for exploring the role of technological advancements and automation in reshaping the demand for generic skills and their associated wage premiums. As technology continues to transform the labour market, understanding how skills like problem-solving, numeracy, and computer complexity evolve in relevance and value can inform strategies for workforce development and adaptation.

Furthermore, considering the findings related to graduates and their choice of study subjects, future research could investigate the effectiveness of educational curricula in fostering generic skill development among students. By examining curriculum designs, teaching methodologies, and experiential learning opportunities, researchers can identify best practices for equipping graduates with the skills demanded by employers and valued in the labour market.

4.7. Appendix

In total, there are four principal components which possess an eigenvalue greater than one. The estimations for the main regression used the principal component with the largest eigenvalue since the component significantly captured the most variance in the dataset.

However, as Table 4.7. illustrates, there are three other principal components which has an eigenvalue greater than one. There are three alternative principal components which the eigenvalue exceeding one. Table 4.8 presents the outcome of a z-test and the p-value associated for the three components to examine if communication skill returns are significantly different when compared to the returns generated by the largest principal component. Specifically, Communication skills 2, 3, and 4 denotes the second, third, and fourth highest principal components.

Table A.4.1.1

Principal components and the eigenvalues of communication skills

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.59	3.39	0.38	0.38
Comp2	1.20	0.08	0.10	0.48
Comp3	1.12	0.10	0.09	0.58
Comp4	1.02	0.35	0.09	0.66
Comp5	0.67	0.06	0.06	0.72
Comp6	0.61	0.04	0.05	0.77
Comp7	0.57	0.03	0.05	0.82
Comp8	0.54	0.07	0.05	0.86
Comp9	0.47	0.04	0.04	0.90
Comp10	0.43	0.05	0.04	0.94
Comp11	0.39	0.01	0.03	0.97
Comp12	0.37	.	0.03	1.00

Note: Table 4.7 shows there are four components that have eigenvalue greater than 1. The largest eigenvalue is significantly greater than the component 2, 3 and 4.

From Table 4.8, it shows that other than communication skills at the 10th quantile, all the other z-values are negative. Since the z-values are negative, this indicates that the communication wage premium lies below the mean compared to the communication wage premium from the estimation used in the original estimation i.e. Figure 4.6.3 and Figure 4.7.3 In terms of the results for communication skills 4, the p-values for the z-values are all significant to the 1% level. Other than the 10th quantile, the other two communication skills also show highly significant results. This confirms that when using the three alternative principal components, the wage premium of communication skills is significantly lower compared to the principal component with the highest eigenvalue.

As a result of the consistent negative z-value with the three communication skills, it can be argued that the principal components with a lower eigenvalue should be used instead of the component that possess the highest eigenvalue. The results produced by the highest eigenvalue may have overestimated the wage premium for communication skills. In this circumstance, balancing the information retention and dimensionality is important when choosing the principal component for the main estimation. Retaining too many components may lead to overfitting, while retaining too few may result in information loss. As an overall, the first principal component is chosen since the eigenvalue is significantly greater than its counterparts, thus it substantially captured the largest amount of variability in the data.

Table A.4.2.1

Z-value and P-value of communication skills when choosing various eigenvectors

Variables	**Q10**		**Q25**		**Q50**		**Q75**		**Q90**	
Generic Skills	**Z-Value**	**P-Value**	**Z-Value**	**P-Value**	**Z-Value**	**P-Value**	**Z-Value**	**P-Value**	**Z-Value**	**P-Value**
Communication Skills 2	-1.24	0.11	-4.00	0.00	-6.41	0.00	-7.58	0.00	-3.88	0.00
Communication Skills 3	0.59	0.28	-2.17	0.02	-5.85	0.00	-7.55	0.00	-3.55	0.00
Communication Skills 4	-4.77	0.00	-4.77	0.00	-9.53	0.00	-8.37	0.00	-3.83	0.00

Note: Table 4.8 shows that overall, the communication skill return is significantly lower than the original estimation when using any of the three alternative principal components i.e. component 2, 3 or 4. Q10 – Q90 denotes the quantile of interest.

Table A.4.3.1

Estimates of the wage premium for males: Control variables included

Distributional Statistic	10th Quantile		25th Quantile		50th Quantile		75th Quantile		90th Quantile	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age	0.024***	0.004	0.022***	0.005	0.026***	0.006	0.024***	0.005	0.023***	0.004
Age Squared (x1000)	-0.021***	0.000	-0.003***	0.000	-0.00***	0.000	-0.023***	0.000	-0.003***	0.000
Tenure Years	0.003***	0.000	0.004***	0.000	0.008***	0.000	0.006*	0.000	0.004	0.000
Tenure Years Squared (x1000)	-0.006***	0.000	-0.007***	0.000	-0.006***	0.000	-0.003	0.000	-0.002	0.000
Supervisor or Manager										
Yes, Supervise other Employees	-	-	-	-	-	-	-	-	-	-
Yes, have Managerial Duties	0.006	0.020	0.006	0.026	0.124***	0.029	0.158***	0.026	0.100***	0.020
No, neither	-0.02+	0.015	-0.077***	0.021	-0.078***	0.022	-0.041**	0.020	-0.022	0.015
Full or Part Time										
Full-Time	-	-	-	-	-	-	-	-	-	-
Part-Time	-0.184***	0.024	-0.215***	0.032	-0.039	0.034	0.058	0.031	0.043	0.023
Highest Education Level										
No Qualifications	-	-	-	-	-	-	-	-	-	-
GCSE D - G	0.055***	0.026	0.048	0.035	0.133***	0.038	0.012	0.035	-0.027	0.026

GCSE A* - C	0.067**	0.025	0.087***	0.034	0.184***	0.036	0.053	0.033	-0.016	0.025
A Levels	0.065**	0.029	0.136***	0.038	0.235***	0.041	0.144***	0.037	0.072***	0.028
University Diploma	0.082**	0.036	0.104**	0.048	0.198***	0.052	0.107**	0.047	0.014	0.036
Undergraduate	0.074**	0.032	0.167***	0.042	0.237***	0.046	0.129***	0.042	0.076	0.032
Master or PhD	0.069*	0.038	0.109*	0.051	0.259***	0.055	0.292***	0.050	0.249***	0.038
Other	0.090***	0.033	0.125***	0.044	0.095**	0.048	0.026	0.044	0.018	0.033
Occupation Level										
Higher Managerial	-	-	-	-	-	-	-	-	-	-
Lower Managerial	0.058***	0.019	0.041	0.025	0.065**	0.027	-0.054**	0.025	-0.035	0.019
Intermediate Occupations	0.056**	0.023	0.033	0.031	-0.037	0.034	-0.056*	0.031	-0.044*	0.023
Small Employers	0.066	0.035	0.038	0.047	0.044	0.051	-0.029	0.046	-0.076**	0.035
Lower Supervisory	0.027	0.025	0.017	0.033	-0.029	0.036	-0.064*	0.032	-0.037	0.025
Semi-Routine	-0.036	0.023	-0.059*	0.031	-0.025	0.033	-0.067**	0.030	-0.048	0.023
Routine	0.004	0.024	-0.056	0.032	-0.020	0.035	-0.048	0.032	-0.029	0.024
Private or Public Sector										
Private	-	-	-	-	-	-	-	-	-	-
Public	0.013	0.015	0.035	0.020	-0.013	0.022	-0.037	0.020	-0.051***	0.015
Non-Profit	0.025	0.036	-0.057	0.047	-0.046	0.051	-0.106**	0.047	-0.098***	0.035
No Answer	-0.054	0.038	-0.040	0.050	0.028	0.054	-0.084	0.049	-0.026	0.037
Firm Size										
Just 1	-	-	-	-	-	-	-	-	-	-

<25	-0.026	0.038	-0.056	0.051	-0.016	0.055	-0.071	0.050	-0.067	0.038
25-99	-0.027	0.039	-0.057	0.052	0.035	0.056	-0.063	0.051	-0.089**	0.039
100-499	0.018	0.039	-0.045	0.053	0.010	0.057	-0.057	0.052	-0.098**	0.039
500+	-0.023	0.041	-0.056	0.054	0.094	0.059	0.026	0.053	0.026	0.040
Marital Status										
Married	-	-	-	-	-	-	-	-	-	-
Living Together as a Couple	-0.014	0.017	-0.055**	0.023	-0.068***	0.025	-0.076***	0.023	-0.055***	0.017
Single	-0.067***	0.017	-0.116***	0.023	-0.146***	0.025	-0.094***	0.023	-0.037**	0.017
Widowed	0.048	0.092	0.194	0.122	0.067	0.132	0.015	0.120	0.036	0.091
Separated or Divorced	-0.020	0.026	0.023	0.034	-0.038	0.037	-0.112***	0.034	-0.038	0.025
Ethnicity										
White British	-	-	-	-	-	-	-	-	-	-
Black - Caribbean	0.032	0.055	0.162**	0.073	0.065	0.079	-0.114	0.072	-0.071	0.054
Black - African	0.035	0.051	-0.114*	0.069	-0.136*	0.074	-0.067	0.068	-0.066	0.051
Black - Other	0.127	0.123	0.306*	0.164	0.167	0.176	0.019	0.161	-0.098	0.121
Indian	0.036	0.038	-0.048	0.050	-0.034	0.054	-0.045	0.049	0.008	0.037
Pakistani	-0.018	0.053	-0.037	0.070	0.006	0.076	-0.057	0.069	-0.059	0.052
Bangladeshi	0.267***	0.092	0.044	0.123	0.094	0.132	0.005	0.120	0.055	0.091
Chinese	-0.029	0.083	0.068	0.111	0.318***	0.119	0.309***	0.109	0.064	0.082
Other	-0.027	0.038	-0.072	0.051	0.012	0.055	-0.155***	0.050	-0.107***	0.038
Region of Workplace										

East Midlands	-	-	-	-	-	-	-	-	-	-
North East	0.047	0.037	0.078	0.049	0.107*	0.053	0.034	0.048	-0.024	0.036
North West	0.035	0.031	-0.024	0.041	0.026	0.045	-0.037	0.041	0.011	0.031
Yorkshire and the Humber	0.049	0.030	0.022	0.040	0.044	0.043	0.029	0.040	-0.013	0.030
West Midlands	0.046	0.031	0.003	0.041	0.011	0.044	0.028	0.040	0.018	0.030
East of England	0.074**	0.030	-0.024	0.040	0.046	0.043	0.055	0.039	0.057	0.030
London	0.072**	0.031	0.102**	0.041	0.159***	0.045	0.162***	0.041	0.098***	0.031
South East	0.071**	0.029	0.033	0.039	0.037	0.042	0.094**	0.038	0.079**	0.029
South West	0.043	0.032	-0.014	0.042	0.065	0.046	0.056	0.042	0.006	0.031
Wales	0.047	0.028	0.036	0.037	0.044	0.040	0.019	0.036	-0.011	0.027
Scottish Lowlands	0.079**	0.033	0.057	0.044	0.071	0.048	0.067	0.044	-0.026	0.033
Generic Skills										
Literature Skills	0.008	0.008	0.016	0.011	0.031***	0.012	0.036***	0.011	0.004	0.008
Numeracy Skills	0.014	0.007	0.007	0.010	-0.012	0.011	-0.017	0.010	0.016	0.007
Communication Skills	0.006	0.008	0.018	0.011	0.036***	0.011	0.029*	0.010	0.018*	0.008
Problem-Solving Skills	0.027***	0.007	0.045***	0.010	0.044***	0.010	0.024**	0.009	0.015	0.007
Physical Skills	-0.026**	0.007	-0.041***	0.010	-0.088***	0.010	-0.086***	0.009	-0.042***	0.007
Computer Complexity	0.031***	0.007	0.063***	0.009	0.084***	0.010	0.045***	0.009	0.014	0.007
Constant	**0.347*** **	**0.101**	**-0.021**	**0.135**	**-0.475*** **	**0.146**	**-0.396*** **	**0.133**	**-0.187***	**0.100**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. Q10 – Q90 denotes the quantile of interest. The dash lines (-) signifies the base category.

Table A.4.3.2

Estimates of the wage premium for females: Control variables included

Distributional Statistic	10th Quantile		25th Quantile		50th Quantile		75th Quantile		90th Quantile	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age	0.017***	0.004	0.015**	0.005	0.017***	0.005	0.025**	0.005	0.016***	0.004
Age Squared (x1000)	-0.009**	0.000	-0.006*	0.000	-0.007**	0.000	-0.009***	0.000	-0.008***	0.000
Tenure Years	0.007***	0.000	0.008***	0.000	0.005***	0.000	0.007	0.000	0.007**	0.000
Tenure Years Squared (x1000)	-0.004**	0.000	-0.009***	0.000	-0.007	0.000	-0.008**	0.000	-0.006***	0.000
Supervisor or Manager										
Yes, Supervise other Employees	-	-	-	-	-	-	-	-	-	-
Yes, have Managerial Duties	0.005	0.020	0.017	0.025	0.131***	0.027	0.145***	0.025	0.076***	0.019
No, neither	-0.038*	0.015	-0.054***	0.019	-0.060***	0.020	-0.054**	0.019	-0.017	0.015
Full or Part Time										
Full-Time	-	-	-	-	-	-	-	-	-	-
Part-Time	-0.047***	0.012	-0.067***	0.016	-0.081***	0.017	-0.028	0.016	0.016	0.012
Highest Education Level										
No Qualifications	-	-	-	-	-	-	-	-	-	-
GCSE D - G	0.108***	0.028	0.129***	0.036	0.070***	0.038	-0.025	0.035	-0.029	0.027

GCSE A* - C	0.115***	0.026	0.204***	0.033	0.144*	0.035	0.018	0.033	0.004	0.025
A Levels	0.167***	0.029	0.278***	0.037	0.226***	0.039	0.117***	0.036	0.062**	0.028
University Diploma	0.136***	0.038	0.235***	0.049	0.279***	0.052	0.115**	0.048	0.020	0.038
Undergraduate	0.142***	0.031	0.266***	0.040	0.304***	0.042	0.206***	0.039	0.071**	0.030
Master or PhD	0.128***	0.039	0.247***	0.051	0.287***	0.054	0.247***	0.050	0.183***	0.039
Other	0.090**	0.044	0.095	0.057	0.043	0.061	0.037	0.056	-0.012	0.044
Occupation Level										
Higher Managerial	-	-	-	-	-	-	-	-	-	-
Lower Managerial	0.065***	0.019	0.066***	0.024	0.072***	0.026	-0.032	0.024	-0.062***	0.018
Intermediate Occupations	0.096***	0.021	0.077***	0.027	-0.064**	0.029	-0.135***	0.027	-0.091	0.021
Small Employers	0.074**	0.033	0.046	0.042	0.090*	0.045	-0.026	0.041	-0.030	0.032
Lower Supervisory	0.057**	0.028	0.059	0.036	-0.043	0.039	-0.187***	0.036	-0.134	0.028
Semi-Routine	0.049*	0.022	-0.032	0.028	-0.134***	0.030	-0.118***	0.028	-0.073	0.021
Routine	0.084***	0.025	0.001	0.032	0.031	0.034	-0.033	0.031	-0.047	0.024
Private or Public Sector										
Private	-	-	-	-	-	-	-	-	-	-
Public	-0.018	0.013	0.026	0.017	0.024	0.018	-0.037	0.017	-0.035**	0.013
Non-Profit	-0.016	0.028	-0.025	0.036	0.017	0.038	-0.086**	0.035	-0.041	0.027
No Answer	-0.014	0.032	0.063	0.042	-0.026	0.044	-0.048	0.041	-0.031	0.032
Firm Size										
Just 1	-	-	-	-	-	-	-	-	-	-

<25	0.063	0.036	0.044	0.046	-0.021	0.049	-0.036	0.045	-0.044	0.035
25-99	0.056	0.037	0.076	0.048	0.034	0.051	-0.019	0.047	-0.048	0.037
100-499	0.047	0.038	0.044	0.048	0.006	0.052	-0.035	0.048	-0.042	0.037
500+	0.065	0.038	0.062	0.050	0.017	0.053	-0.014	0.049	-0.016	0.038
Marital Status										
Married	-	-	-	-	-	-	-	-	-	-
Living Together as a Couple	0.021	0.017	0.020	0.022	-0.056*	0.024	-0.066***	0.022	-0.020	0.017
Single	-0.026	0.016	-0.054***	0.021	-0.077***	0.022	-0.065***	0.020	-0.037*	0.016
Widowed	0.007	0.040	-0.058	0.052	-0.078	0.055	-0.054	0.051	0.009	0.040
Separated or Divorced	-0.026	0.017	-0.034	0.022	-0.070***	0.024	-0.057**	0.022	-0.034*	0.017
Ethnicity										
White British	-	-	-	-	-	-	-	-	-	-
Black - Caribbean	0.008	0.067	0.065	0.087	0.104	0.092	0.126	0.085	0.006	0.066
Black - African	-0.065	0.043	-0.024	0.055	-0.146**	0.059	-0.167***	0.054	-0.097**	0.042
Black - Other	-0.034	0.106	0.096	0.137	0.327**	0.146	0.166	0.134	0.179	0.104
Indian	0.016	0.048	0.047	0.063	0.006	0.067	0.048	0.061	-0.017	0.048
Pakistani	0.137*	0.073	-0.026	0.094	0.019	0.100	0.005	0.093	0.094	0.072
Bangladeshi	0.098	0.141	0.157	0.182	0.037	0.193	-0.051	0.179	-0.266*	0.139
Chinese	0.136	0.085	0.135	0.110	0.015	0.117	0.013	0.108	0.204**	0.084
Other	0.039	0.035	0.011	0.046	-0.094*	0.049	-0.075*	0.045	-0.045	0.035
Region of Workplace										

East Midlands	-	-	-	-	-	-	-	-	-	-
North East	-0.036	0.034	-0.015	0.043	0.102	0.046	0.076	0.043	0.000	0.033
North West	-0.034	0.027	-0.038	0.035	0.036	0.037	0.069	0.034	0.007	0.027
Yorkshire and the Humber	-0.027	0.028	-0.054	0.036	0.007	0.038	0.027	0.035	0.006	0.027
West Midlands	0.006	0.028	0.014	0.037	0.044	0.039	0.058	0.036	-0.017	0.028
East of England	0.045	0.028	0.043	0.036	0.095	0.038	0.136***	0.035	0.028	0.028
London	0.047	0.032	0.033	0.041	0.118***	0.043	0.247***	0.040	0.147***	0.031
South East	0.046	0.027	0.037	0.035	0.076*	0.037	0.104***	0.034	0.054*	0.027
South West	0.021	0.029	0.014	0.037	0.004	0.039	0.077*	0.036	0.026	0.028
Wales	-0.010	0.025	-0.030	0.032	0.021	0.034	0.039	0.031	0.004	0.024
Scottish Lowlands	-0.023	0.030	-0.044	0.039	0.014	0.041	0.071*	0.038	-0.013	0.030
Generic Skills										
Literature Skills	0.027***	0.007	0.032***	0.009	0.057***	0.010	0.048***	0.009	0.012**	0.007
Numeracy Skills	-0.028***	0.006	-0.034***	0.008	-0.027***	0.009	-0.026**	0.008	0.004	0.006
Communication Skills	0.009	0.007	0.038***	0.009	0.066***	0.010	0.077***	0.009	0.034***	0.007
Problem-Solving Skills	0.007	0.006	0.026**	0.007	0.009	0.008	0.009	0.007	0.011**	0.006
Physical Skills	-0.019**	0.006	-0.064***	0.008	-0.074***	0.009	-0.054***	0.008	-0.047***	0.006
Computer Complexity	0.067***	0.007	0.100***	0.009	0.084***	0.010	0.045***	0.009	0.019	0.007
Constant	**0.285*** **	**0.097**	**-0.112**	**0.126**	**-0.241***	**0.134**	**-0.253*** **	**0.123**	**-0.108**	**0.096**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. Q10 – Q90 denotes the quantile of interest. The dash lines (-) signifies the base category.

Table A.4.3.3

Estimates of the wage premium for White: Control variables included

Distributional Statistic	10th Quantile		25th Quantile		50th Quantile		75th Quantile		90th Quantile	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age	0.016***	0.003	0.024***	0.004	0.025***	0.004	0.026***	0.004	0.014***	0.003
Age Squared	-0.007***	0.000	-0.006***	0.000	-0.007***	0.000	-0.007***	0.000	-0.006***	0.000
Tenure Years	0.006***	0.000	0.009***	0.000	0.006***	0.000	0.006	0.000	0.008	0.000
Tenure Years Squared	-0.002***	0.000	-0.004***	0.000	-0.004***	0.000	-0.007	0.000	-0.002	0.000
Supervisor or Manager										
Yes, Supervise other Employees	-	-	-	-	-	-	-	-	-	-
Yes, have Managerial Duties	0.008	0.0158	0.001	0.019	0.126	0.021	0.140	0.019	0.084	0.014
No, neither	-0.016	0.0114	-0.056***	0.015	-0.091***	0.016	-0.055***	0.015	-0.025**	0.011
Full or Part Time										
Full-Time	-	-	-	-	-	-	-	-	-	-
Part-Time	-0.085***	0.011	-0.179***	0.014	-0.105***	0.015	-0.037**	0.014	0.009	0.011
Highest Education Level										

No Qualifications	-	-	-	-	-	-	-	-	-	-
GCSE D - G	0.095***	0.020	0.094***	0.026	0.095***	0.028	-0.016	0.026	-0.025	0.020
GCSE A* - C	0.104***	0.019	0.151***	0.024	0.142***	0.027	0.036	0.024	0.008	0.019
A Levels	0.146***	0.021	0.203***	0.027	0.213***	0.030	0.125***	0.027	0.079***	0.021
University Diploma	0.107***	0.028	0.175***	0.036	0.232***	0.040	0.114***	0.036	0.024	0.028
Undergraduate	0.145***	0.024	0.227***	0.030	0.274***	0.033	0.151***	0.030	0.055**	0.023
Master or PhD	0.126***	0.030	0.186***	0.039	0.237***	0.043	0.272***	0.039	0.233***	0.029
Other	0.142***	0.028	0.218***	0.036	0.138***	0.040	0.057***	0.036	0.004	0.027
Occupation Level										
Higher Managerial	-	-	-	-	-	-	-	-	-	-
Lower Managerial	0.062***	0.014	0.067***	0.018	0.076***	0.020	-0.045**	0.018	-0.053***	0.014
Intermediate Occupations	0.073***	0.016	0.045**	0.021	-0.055**	0.023	-0.106***	0.021	-0.071***	0.016
Small Employers	0.075***	0.025	0.038	0.032	0.076*	0.036	-0.037	0.032	-0.047	0.025
Lower Supervisory	0.062***	0.020	0.035	0.025	-0.027	0.028	-0.086***	0.025	-0.056***	0.019
Semi-Routine	0.001	0.017	-0.067***	0.021	-0.084***	0.024	-0.078***	0.021	-0.057***	0.016
Routine	0.063***	0.018	-0.014	0.024	0.026	0.026	-0.045*	0.024	-0.038	0.018
Private or Public Sector										
Private	-	-	-	-	-	-	-	-	-	-
Public	-0.026	0.010	0.025*	0.013	0.008	0.015	-0.032***	0.013	-0.041***	0.010

Non-Profit	-0.013	0.022	0.007	0.029	-0.026	0.032	-0.094***	0.029	-0.105***	0.022
No Answer	-0.037	0.026	-0.015	0.033	-0.027	0.036	-0.070**	0.033	-0.046*	0.025
Firm Size										
Just 1	-	-	-	-	-	-	-	-	-	-
<25	0.036	0.027	-0.026	0.035	-0.024	0.039	-0.054	0.035	-0.029	0.027
25-99	0.038	0.029	0.005	0.037	0.018	0.040	-0.047	0.037	-0.036	0.028
100-499	0.047	0.029	0.012	0.037	0.005	0.041	-0.054	0.037	-0.042	0.028
500+	0.034	0.030	-0.010	0.038	0.023	0.042	-0.026	0.038	0.023	0.029
Marital Status										
Married	-	-	-	-	-	-	-	-	-	-
Living Together as a Couple	0.014	0.013	-0.019	0.016	-0.058***	0.018	-0.064***	0.016	-0.047***	0.012
Single	-0.028*	0.012	-0.075***	0.016	-0.095***	0.017	-0.061***	0.016	-0.058***	0.012
Widowed	-0.017	0.038	-0.044	0.048	-0.129**	0.053	-0.056	0.048	-0.017	0.037
Separated or Divorced	-0.026	0.015	-0.061***	0.019	-0.085***	0.021	-0.073***	0.019	-0.042***	0.015
Region of Workplace										
East Midlands	-	-	-	-	-	-	-	-	-	-
North East	-0.013	0.026	0.019	0.033	0.076*	0.036	0.044	0.033	-0.019	0.025
North West	-0.017	0.021	-0.037	0.027	0.002	0.030	0.026	0.027	-0.047*	0.021

Yorkshire and the Humber	0.005	0.021	-0.020	0.027	-0.037	0.030	0.024	0.027	0.007	0.021
West Midlands	0.016	0.022	0.034	0.028	0.006	0.031	0.017	0.028	-0.019	0.021
East of England	0.048*	0.021	0.056**	0.028	0.068**	0.030	0.094***	0.028	0.038	0.021
London	0.029	0.025	0.048	0.032	0.132***	0.036	0.216***	0.032	0.165***	0.025
South East	0.034	0.021	0.067**	0.027	0.033	0.029	0.089***	0.027	0.056	0.020
South West	0.032	0.022	0.039	0.028	0.014	0.031	0.078**	0.028	0.024	0.022
Wales	0.004	0.019	0.018	0.025	-0.021	0.027	0.012	0.024	-0.036	0.019
Scottish Lowlands	0.002	0.023	0.002	0.030	0.010	0.033	0.060**	0.030	-0.036	0.023
Generic Skills										
Literature Skills	0.024***	0.006	0.021***	0.007	0.034***	0.008	0.028***	0.007	0.004	0.005
Numeracy Skills	-0.015*	0.005	0.008	0.006	-0.017	0.007	-0.014	0.006	0.016	0.005
Communication Skills	0.008	0.006	0.027***	0.007	0.056***	0.008	0.076***	0.007	0.034***	0.006
Problem-Solving Skills	0.007*	0.005	0.026***	0.006	0.038**	0.007	0.019**	0.006	0.018***	0.005
Physical Skills	-0.016***	0.005	-0.049***	0.006	-0.062***	0.007	-0.062***	0.006	-0.042***	0.005
Computer Complexity	0.049***	0.005	0.072***	0.007	0.074***	0.007	0.041***	0.007	0.010**	0.005
Constant	**0.256*** **	**0.073**	**-0.124**	**0.094**	**-0.407*** **	**0.103**	**-0.298*** **	**0.094**	**-0.126***	**0.072**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. Q10 – Q90 denotes the quantile of interest. The dash lines (-) signifies the base category.

Table A.4.3.4

Estimates of the wage premium for non-White: Control variables included

Distributional Statistic	10th Quantile		25th Quantile		50th Quantile		75th Quantile		90th Quantile	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age	-0.018	0.011	0.009	0.014	0.022	0.015	0.020	0.014	0.012	0.010
Age Squared	-0.005	0.000	-0.007	0.000	-0.004	0.000	-0.005	0.000	-0.004	0.000
Tenure Years	0.004**	0.000	0.005***	0.001	0.003	0.001	0.004	0.001	0.003	0.000
Tenure Years Squared	-0.006	0.000	-0.004**	0.000	-0.008	0.000	-0.005	0.000	-0.008	0.000
Supervisor or Manager										
Yes, Supervise other Employees	-	-	-	-	-	-	-	-	-	-
Yes, have Managerial Duties	-0.031	0.051	-0.046	0.069	0.075	0.070	0.088	0.066	0.218***	0.050
No, neither	-0.010	0.036	-0.083	0.048	-0.039	0.049	-0.147***	0.046	-0.026	0.035
Full or Part Time										
Full-Time	-	-	-	-	-	-	-	-	-	-
Part-Time	-0.075**	0.035	-0.157**	0.047	-0.152***	0.048	-0.023	0.045	0.007	0.035
Highest Education Level										

No Qualifications	-	-	-	-	-	-	-	-	-	-
GCSE D - G	-0.108	0.065	-0.105	0.088	-0.056	0.090	-0.078	0.084	0.017	0.064
GCSE A* - C	-0.086	0.059	-0.134	0.080	0.083	0.081	-0.027	0.076	0.024	0.058
A Levels	-0.065	0.065	-0.102	0.089	0.042	0.090	-0.072	0.085	0.046	0.065
University Diploma	0.023	0.076	-0.091	0.103	0.134	0.105	0.004	0.099	-0.042	0.075
Undergraduate	-0.064	0.063	-0.073	0.085	0.173*	0.087	0.116	0.081	0.113*	0.062
Master or PhD	-0.136*	0.076	-0.087	0.103	0.138	0.105	0.177*	0.099	0.084	0.075
Other	-0.237***	0.085	-0.318***	0.115	-0.049	0.117	-0.068	0.110	-0.033	0.084
Occupation Level										
Higher Managerial	-	-	-	-	-	-	-	-	-	-
Lower Managerial	0.030	0.047	0.037	0.064	-0.107	0.066	-0.126*	0.061	-0.157***	0.047
Intermediate Occupations	0.037	0.055	0.054	0.075	-0.284***	0.076	-0.218***	0.072	-0.124**	0.055
Small Employers	-0.111	0.078	-0.104	0.106	-0.322***	0.108	-0.075	0.101	-0.058	0.077
Lower Supervisory	0.042	0.064	-0.081	0.086	-0.233***	0.088	-0.193**	0.083	-0.155**	0.063
Semi-Routine	0.013	0.053	-0.042	0.071	-0.227***	0.073	-0.184***	0.068	-0.096*	0.052
Routine	0.088	0.055	0.063	0.075	-0.143*	0.076	-0.142**	0.071	-0.087	0.054
Private or Public Sector										
Private	-	-	-	-	-	-	-	-	-	-
Public	0.005	0.032	0.037	0.044	0.029	0.045	0.044	0.042	0.004	0.032

Non-Profit	-0.017	0.113	0.165	0.152	-0.26**	0.156	-0.221	0.146	-0.166	0.111
No Answer	0.094	0.086	0.26/**	0.116	0.387***	0.119	0.030	0.111	-0.077	0.085
Firm Size										
Just 1	-	-	-	-	-	-	-	-	-	-
<25	0.147	0.091	0.326***	0.123	0.156	0.125	-0.046	0.117	-0.105	0.090
25-99	0.209**	0.093	0.345***	0.125	0.199	0.128	-0.034	0.120	-0.147	0.091
100-499	0.127	0.093	0.276**	0.125	0.205	0.128	-0.028	0.120	-0.136	0.092
500+	0.181*	0.095	0.335***	0.129	0.264**	0.132	0.057	0.124	0.031	0.094
Marital Status										
Married	-	-	-	-	-	-	-	-	-	-
Living Together as a Couple	-0.013	0.046	0.014	0.063	-0.024	0.064	-0.127**	0.060	-0.075	0.046
Single	-0.122***	0.040	-0.096*	0.055	-0.067	0.056	-0.075	0.052	-0.054	0.040
Widowed	-0.123	0.151	-0.105	0.204	0.236	0.208	0.073	0.195	0.188	0.149
Separated or Divorced	-0.047	0.052	-0.021	0.071	-0.089	0.072	-0.012	0.068	-0.042	0.052
Region of Workplace										
East Midlands	-	-	-	-	-	-	-	-	-	-
North East	0.052	0.142	0.284	0.192	0.119	0.196	-0.068	0.184	0.063	0.140
North West	-0.025	0.077	0.046	0.104	0.055	0.106	0.029	0.099	0.055	0.076

Yorkshire and the Humber	-0.107	0.075	-0.044	0.101	0.284***	0.104	-0.174*	0.097	-0.084	0.074
West Midlands	-0.016	0.074	0.076	0.100	0.218**	0.102	-0.012	0.095	-0.087	0.073
East of England	0.108	0.072	0.121	0.098	0.194*	0.100	0.081	0.093	-0.016	0.071
London	0.035	0.063	0.127	0.085	0.241***	0.087	0.093	0.081	0.039	0.062
South East	-0.021	0.071	0.024	0.096	0.073	0.098	0.095	0.092	0.033	0.070
South West	-0.122	0.087	-0.018	0.117	0.132	0.120	0.152	0.112	-0.012	0.086
Wales	-0.197**	0.083	-0.049	0.112	0.130	0.115	-0.041	0.107	-0.052	0.082
Scottish Lowlands	0.048	0.101	0.122	0.137	0.156	0.140	0.133	0.131	0.012	0.100
Generic Skills										
Literature Skills	0.022	0.018	0.034	0.024	0.024	0.025	0.035	0.023	0.032	0.018
Numeracy Skills	0.004	0.016	0.002	0.021	-0.027	0.022	-0.024	0.020	-0.023	0.016
Communication Skills	0.021	0.017	0.051	0.024	0.094***	0.024	0.056**	0.023	0.024	0.017
Problem-Solving Skills	-0.013	0.016	-0.023	0.022	0.013	0.023	-0.019	0.021	-0.025	0.016
Physical Skills	-0.048**	0.016	-0.066***	0.022	-0.142***	0.022	-0.107***	0.021	-0.041***	0.016
Computer Complexity	0.043***	0.016	0.088***	0.021	0.070***	0.022	0.048**	0.020	0.010	0.016
Constant	**0.974*** **	**0.253**	**0.321**	**0.343**	**-0.378**	**0.350**	**0.026**	**0.328**	**0.065**	**0.250**

Note: Significance at the 1%, 5% and 10% levels denoted by ***, ** and * respectively. The dependent variable is real earnings in 2010 prices. Q10 – Q90 denotes the quantile of interest. The dash lines (-) signifies the base category.

Chapter 5

Conclusion

5.1. Summary

This thesis brings forward a comprehensive examination of the complex dynamics between skills, migration, wage outcomes, and productivity constraints within the UK labour market, offering actionable insights for the UK's Industrial Strategy. The findings highlight the need for a skills-focused approach to address regional disparities, resolve skill mismatches, and prepare for evolving market demands. By delving into the intricate relationships between graduate mobility, job mismatches, and the economic returns on generic skills, this research underscores critical areas where targeted policy interventions could enhance both economic resilience and labour market equity across the UK.

One key insight concerns graduate migration patterns and the associated wage outcomes. Graduates who migrate for education but return to their original regions, along with those who remain in their hometowns throughout their education and careers, tend to encounter significant wage penalties. This trend suggests that the UK's industrial and regional policies must more effectively address the challenges of "brain drain" in regions outside of London and the South East. Although universities play a critical role in enhancing local human capital, the limited availability of high-paying, high-skilled jobs lead many graduates to seek better opportunities elsewhere, often in urban centres. This outmigration exacerbates regional economic disparities, hindering local innovation, productivity, and growth in less affluent regions.

For the Industrial Strategy, these findings indicate that regional economic development policies should prioritize job creation and competitive wage structures to attract and retain highly skilled graduates. Policy solutions might include expanding infrastructure, incentivizing business investment outside major urban hubs, and establishing partnerships between local governments, universities, and private enterprises. Such initiatives could foster a more dynamic and self-sustaining economic landscape in regions that traditionally suffer from outmigration, helping to create more balanced economic growth and alleviate regional skill shortages.

Further insights are gained from the analysis of job mismatches and their impact on wage outcomes. Educational mismatches, both vertical (where individuals are over- or under-qualified for their roles) and horizontal (where job roles do not align with individuals' fields of study), were found to have notable effects on wages and job satisfaction. Vertical mismatches, in particular, are detrimental to both individual and economic productivity, as overqualified workers frequently face underutilization of their skills, leading to job dissatisfaction and lower productivity. These issues are even more pronounced for ethnic minorities, who face additional structural barriers that limit their access to roles that align with their qualifications, further contributing to wage disparities.

Furthermore, horizontal mismatches also pose challenges for career progression and economic mobility, as workers employed in fields unrelated to their studies often lack the specialized skills necessary for advancement, which limits wage growth. Addressing these mismatches would require educational reforms aimed at aligning programs with the evolving needs of the labour market. Greater collaboration between universities and industries, targeted apprenticeships, work placements, and region-specific skill programs could reduce job mismatches, particularly for minority groups disproportionately affected by these barriers. By enhancing job alignment, such policies could lower wage penalties, bolster productivity, and address skill shortages within the workforce.

The study also brings to light the significant wage premiums associated with generic skills, particularly advanced digital and computer skills, which are increasingly valued across all industries. The research shows that digital skills yield higher economic returns compared to many other generic skills, reflecting the growing demand for digital competencies in the modern economy. With technological advancements driving increased demand for such skills, policy support for digital and computer skills development becomes crucial for maintaining the UK’s competitiveness in a globalised economy. Targeted investments in digital skills training across all levels of education—from primary and secondary schooling to adult reskilling programs—could be essential to equipping the workforce with the competencies needed to succeed in a technology-driven labour market.

Moreover, wage returns for digital skills vary by demographic factors such as gender and ethnicity, suggesting that disparities in access to these training opportunities contribute to broader wage gaps. By expanding access to digital skill development, particularly for women and ethnic minorities, policies could reduce these disparities and promote a more inclusive and equitable labour market. Strategic partnerships between technology firms and educational institutions could ensure that digital training programs remain industry-relevant, addressing current skill shortages while also preparing for future demands.

The broader implications of this thesis for the UK's Industrial Strategy revolve around the central role of skills development as a foundation for economic resilience, regional prosperity, and productivity growth. The research supports the notion that a skills-focused approach is essential for resolving the UK's productivity puzzle, which is marked by stagnating productivity growth despite an increasingly educated workforce. Strategic investments in high-demand areas, such as STEM and digital skills, are pivotal for enhancing workforce adaptability, driving innovation, and ultimately fuelling economic growth. Addressing skill mismatches and fostering the retention of high-skilled workers across regions could also support balanced economic development, helping to reduce the country's reliance on a few economically prosperous areas and creating growth opportunities throughout the nation.

Ultimately, the successful implementation of the UK's Industrial Strategy will depend on developing a comprehensive skills policy that is adaptable to regional needs and aligned with future labour market trends. This approach must go beyond addressing immediate skill shortages and also anticipate future shifts in demand, allowing the UK workforce to remain competitive in a rapidly evolving global economy. Prioritizing inclusive skill development, regionally targeted investment, and stronger partnerships between education providers and employers could support a more dynamic, equitable, and resilient economy that benefits all regions and demographics.

Addressing capital-labour complementarities and productivity constraints remains essential to tackling the challenges outlined in this thesis. As industries increasingly rely on advanced technical and digital skills, the demand for skilled labour that complements capital investment has risen. However, the alignment between available skills and market demand is often insufficient, particularly in economically disadvantaged regions, where firms struggle to attract and retain talent, frequently due to "brain drain" migration toward more affluent areas. Investment in generic skills, including digital literacy, problem-solving, and adaptability, has proven effective in raising productivity and wage premiums. Nevertheless, uneven investment and opportunities in skills development across regions perpetuate regional inequalities and a cyclical pattern of underutilized human capital in less prosperous areas.

Additionally, wage premiums related to skill levels vary significantly by ethnicity and gender, presenting further challenges to achieving equitable productivity growth. Groups facing structural disadvantages in skill valuation experience lower returns on skills, which discourages participation in high demand fields and ultimately hampers labour market productivity. To address these barriers, policies that enhance inclusivity in skills training, foster regional retention of skilled labour, and support alignment between skill development and capital investment are critical. A comprehensive approach to capital-labour complementarities—encompassing skill-building, inclusive training, and region-specific retention initiatives—is necessary to overcome current productivity constraints. Such an approach would not only increase productivity and resilience but also support balanced development, ensuring that all regions benefit from technological advancements and access to skilled jobs.

In conclusion, the findings from this thesis underscore the importance of a holistic approach to skills policy within the UK’s Industrial Strategy. By aligning educational programs with labour market demands, investing in high-demand digital and STEM skills, and promoting inclusivity and regional retention, the Industrial Strategy can drive balanced economic growth, reduce regional disparities, and build a resilient and equitable economy that is prepared for future shifts in global demand.

Reference list

Abel, J. R., & Deitz, R. (2009). *Do colleges and universities increase their region's human capital?* (Staff Report No. 401). Federal Reserve Bank of New York.

Abraham, K. G., & Spletzer, J. R. (2009). New evidence on the returns to job skills. *American Economic Review*, *99*(2), 52-57.

Acemoglu, D., & Pischke, J.-S. (1999). Beyond Becker: Training in imperfect labour markets. *The Economic Journal*, 109(453), F112–F142.

Ahlin, L., Andersson, M., & Thulin, P. (2014). Market thickness and the early labour market career of university graduates: an urban advantage?. *Spatial Economic Analysis, 9*(4), 396-419.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Ilr Review*, *30*(2), 175-187.

Alba-Ramirez, A. (1993). Mismatch in the Spanish labor market: Overeducation? *Journal of Human Resources*, *28*(2), 259–278.

Algan, Y., & Cahuc, P. (2009). Civic virtue and labor market institutions. American Economic *Journal: Macroeconomics*, *1*(1), 111-145.

Ammermüller, A. (2005). *Educational opportunities and the role of institutions* (ZEW-Centre for European Economic Research Discussion Paper No. 05-044). ZEW.

Ananiadou, K., Jenkins, A., & Wolf, A. (2004). Basic skills and workplace learning: What do we actually know about their benefits? *Studies in Continuing Education, 26*(2), 289–308.

Ananiadou, K., Jenkins, A., & Wolf, A. (2003). *The benefits to employers of raising workforce basic skills levels: A review of the literature*. National Research and Development Centre for Adult Literacy and Numeracy.

Andrade-Arenas, L., Vasquez-Paragulla, J., & Sotomayor-Beltran, C. (2019, September). On the impact of mathematics in a first-year algorithms course of a computer science program in a university of Lima, Peru. In *2019 International Symposium on Engineering Accreditation and Education (ICACIT)* (pp. 1–4). IEEE.

Appelbaum, E. (2000). *Manufacturing advantage: Why high-performance work systems pay off*. Cornell University Press.

Arevalo, J., Pitknen, S., Gritten, D., & Tahvanainen, L. (2010). Market-relevant competencies for professional foresters in European graduate education. *International Forestry Review*, *12*(3), 200-208.

Arnott, R. (1998). Congestion tolling and urban spatial structure. *Journal of Regional Science*, *38*(3), 601–620.

Arrow, K. J. (1973). Higher education as a filter. *Journal of public economics*, *2*(3), 193-216.

Ashton, D., & Felstead, A. (1998). *Organisational characteristics and skill formation in Britain: Is there a link?* (CLMS Working Paper No. 22). https://leicester.figshare.com/articles/report/Organisational_Characteristics_and_skill_formation_in_Britain_Is_there_a_link_/10099250/files/18205619.pdf

Attewell, P. (1990). What is skill?. *Work and occupations*, *17*(4), 422-448.

Autor, D. (2010). The polarization of job opportunities in the US labor market: Implications

for employment and earnings. *Center for American Progress and The Hamilton Project*, *6*, 11-19.
Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American economic review*, *103*(5), 1553-1597.
Autor, D. H., & Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of labor Economics*, *31*(S1), S59-S96.
Badcock, P. B., Pattison, P. E., & Harris, K. L. (2010). Developing generic skills through university study: A study of arts, science, and engineering in Australia. *Higher Education, 60*(4), 441–458.
Bagguley, P., & Hussain, Y. (2007). *The role of higher education in providing opportunities for South Asian women* (Vol. 2058). Policy Press.
Ballantine*, J. A., & Larres, P. M. (2004). A critical analysis of students' perceptions of the usefulness of the case study method in an advanced management accounting module: the impact of relevant work experience. *Accounting Education*, *13*(2), 171-189.
Barrett, A., & O'Connell, P. J. (2001). Does training generally work? The returns to in-company training. *ILR Review*, *54*(3), 647-662.
Barrett, G., & Staneva, A. (2017). Gender pay gap and returns to cognitive and non-cognitive skills: Evidence from Australia. In *Australian Conference of Economists* (Vol. 6).
Barsky, R., Bound, J., Charles, K. K., & Lupton, J. P. (2002). Accounting for the black–white wealth gap: a nonparametric approach. *Journal of the American statistical Association, 97*(459), 663-673.
Bartel, A. P., & Lichtenberg, F. R. (1987). The skill distribution and competitive trade advantage of high-technology industries. *Advances in Industrial and Labor Relations*, *4*, 161-176.
Bartel, A. P., Ichniowski, C., & Shaw, K. (2003). new technology" and its impact on the jobs of high school educated workers: A look deep inside three manufacturing industries. *Low-Wage America*, 155-94.
Basit, T. N. (1997). 'I Want More Freedom, but Not Too Much': British Muslim girls and the dynamism of family values. *Gender and education, 9*(4), 425-440.
Battistin, E., De Nadai, M., & Sianesi, B. (2014). Misreported schooling, multiple measures and returns to educational qualifications. *Journal of Econometrics, 181*(2), 136-150.
Battu, H., & Sloane, P. J. (2002). To what extent are ethnic minorities in Britain over-educated?. *International Journal of Manpower, 23*(3), 192-208.
Battu, H., Belfield, C. R., & Sloane, P. J. (2000). How well can we measure graduate over-education and its effects?. *National institute economic review, 171*, 82-93.
Baumol, W. J. (1967). Macroeconomics of unbalanced growth: The anatomy of urban crisis. *The American Economic Review, 57*(3), 415–426.
Beaudry, P., & Lewis, E. (2014). Do male-female wage differentials reflect differences in the return to skill? Cross-city evidence from 1980–2000. *American Economic Journal: Applied Economics*, *6*(2), 178-194.
Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science, 9*(3), 226-232.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.
Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.
Becker, G. S. (1975). *Human capital: A theoretical and empirical analysis, with special reference to education* (2nd ed.). University of Chicago Press.
Belfield, C., Britton, J., Buscha, F., Dearden, L., Dickson, M., van der Erve, L., Sibieta, L., Vignoles, A., Walker, I., & Zhu, Y. (2018). *The relative labour market returns to different degrees: Research report*. Institute for Fiscal Studies. https://www.gov.uk/government/publications/undergraduate-degrees-relative-labour-market-returns
Bennett, N., Dunne, E., & Carré, C. (1999). Patterns of core and generic skill provision in higher education. *Higher education, 37*(1), 71-93.
Bergmann, B. R. (1974). Occupational segregation, wages, and profits when employers discriminate by race or sex. *Eastern Economic Journal*, 1(2), 103–110.
Bhopal, K. (2010, May). Gender, identity and experience: Researching marginalised groups. *In Women's Studies International Forum, 33*(3), 188-195. Pergamon.
Bisello, M., Peruffo, E., Fernández-Macías, E., & Rinaldi, R. (2019). *How computerisation is transforming jobs: Evidence from the Eurofound's European Working Conditions Survey* (No. 2019/02). JRC working papers series on Labour, Education and Technology.
Black, S. E., Muller, C., Spitz-Oener, A., He, Z., Hung, K., & Warren, J. R. (2021). The importance of STEM: High school knowledge, skills, and occupations in an era of growing inequality. *Research Policy, 50*(7), Article 104249. https://doi.org/10.1016/j.respol.2021.104249
Blackaby, D. H., Leslie, D. G., Murphy, P. D., & O'Leary, N. C. (1998). The ethnic wage gap and employment differentials in the 1990s: Evidence for Britain. *Economics Letters, 58*(1), 97-103.
Blackaby, D. H., Leslie, D. G., Murphy, P. D., & O'Leary, N. C. (2002). White/ethnic minority earnings and employment differentials in Britain: Evidence from the LFS. *Oxford Economic Papers*, *54*(2), 270-297
Blackaby, D., Leslie, D., Murphy, P., & O'Leary, N. (1999). Unemployment among Britain's ethnic minorities. *Scottish Journal of Political Economy, 44*(1), 182-197.
Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, *8*(4), 436-455.
Boero, G., Nathwani, T., Naylor, R., & Smith, J. (2021). *Graduate earnings premia in the UK: Decline and fall?* Higher Education Statistics Agency.
Bol, T., & Heisig, J. P. (2021). Explaining wage differentials by field of study among higher education graduates: Evidence from a large-scale survey of adult skills. *Social Science Research, 99*, 102594. https://doi.org/10.1016/j.ssresearch.2021.102594
Bond, C. H., Spronken-Smith, R., McLean, A., Smith, N., Frielick, S., Jenkins, M., & Marshall, S. (2017). A framework for enabling graduate outcomes in undergraduate programmes. *Higher Education Research & Development, 36*(1), 43-58.
Bond, R., Charsley, K., & Grundy, S. (2008). Scottish Graduate Migration and Retention: a case study of the University of Edinburgh 2000 cohort. *Scottish Affairs, 63*(1), 31-57.

Borah, B. J., & Basu, A. (2013). Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. *Health economics*, *22*(9), 1052-1070.

Borgen, N. T., Haupt, A., & Wiborg, Ø. N. (2023). Quantile regression estimands and models: revisiting the motherhood wage penalty debate. *European Sociological Review*, *39*(2), 317-331.

Borghans, L., & Ter Weel, B. (2004). Are computer skills the new basic skills? The returns to computer, writing and math skills in Britain. *Labour Economics*, *11*(1), 85-98.

Boudarbat, B., & Chernoff, V. (2012). Education-job match among recent Canadian university graduates. *Applied Economics Letters, 19*(18), 1923-1926.

Boyce, G., Williams, S., Kelly, A., & Yee, H. (2001). Fostering deep and elaborative learning and generic (soft) skill development: the strategic use of case studies in accounting education. *Accounting education*, *10*(1), 37-60.

Boyle, P. (1993). Modelling the relationship between tenure and migration in England and Wales. *Transactions of the Institute of British Geographers*, *18*(4), 359–376.

Bristow, G. I., Pill, M. C., Davies, R., & Drinkwater, S. (2011). *Stay, leave or return? Understanding Welsh graduate mobility* (WISERD Working Paper Series, Vol. 005). Cardiff University. https://orca.cardiff.ac.uk/id/eprint/52711/1/WISERD_WPS_005.pdf

Bristow, G., & Healy, A. (2014). Building resilient regions: Complex adaptive systems and the role of policy intervention. *Raumforschung und Raumordnung, Spatial Research and Planning*, *72*(2), 93-102.

Bristow, S. (2013). *Workforce preparedness of information systems students: Perceptions of students, alumni, and employers* (Graduate Theses and Dissertations). https://scholarworks.uark.edu/etd/1015

Britton, J., Dearden, L., Shephard, N., & Vignoles, A. (2016). *How English domiciled graduate earnings vary with gender, institution attended, subject and socio-economic background.* Institute for Fiscal Studies working paper.

Britton, J., Farquharson, C., Sibieta, L., Tahir, I., & Waltmann, B. (2020). *2020 annual report on education spending in England* (No. R183). Institute for Fiscal Studies.

Brooks, R. (2014). *Contemporary debates in education: An historical perspective*. Routledge.

Brophy, S. (2024). Ethnicity and UK graduate migration: An identity economics approach,. *Journal of Regional Science*, *64*(3), 732-785.

Brynin, M., & Güveli, A. (2012). Understanding the ethnic pay gap in Britain. *Work, employment and society*, *26*(4), 574-587.

Brynin, M., & Longhi, S. (2015). *The effect of occupation on poverty among ethnic minority groups*. Joseph Rowntree Foundation. https://www.jrf.org.uk/report/effect-occupation-poverty-among-ethnic-minority-groups

Card, D., & Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of political Economy, 100*(1), 1-40.

Carliner, G. (1980). Wages, earnings and hours of first, second, and third generation American males. *Economic inquiry, 18*(1), 87-102.

Carrington, E. W. (1993). *The future of education in the Caribbean: Report of the CARICOM Advisory Task Force on Education*. Caribbean Community Secretariat. Layout by The Commonwealth of Learning. Published with assistance from UNESCO.

Castells, M., & Hall, P. (1994). *Technopoles of the world: The making of 21st century industrial complexes* (1st ed.). Routledge.

Catalyst (Organization). (2007). *The double-bind dilemma for women in leadership: Damned if you do, doomed if you don't*. Catalyst.

Catney, G. (2016). The changing geographies of ethnic diversity in England and Wales, 1991–2011. *Population, Space and Place*, *22*(8), 750-765.

Champion, T. (1999). Migration and British cities in the 1990s. *National Institute Economic Review, 170*, 60-77.

Chan, C. K., & Fong, E. T. (2018). Disciplinary differences and implications for the development of generic skills: a study of engineering and business students' perceptions of generic skills. *European Journal of Engineering Education*, *43*(6), 927-949.

Chan, C. K., Zhao, Y., & Luk, L. Y. (2017). A validated and reliable instrument investigating engineering students' perceptions of competency in generic skills. *Journal of Engineering Education*, *106*(2), 299-325.

Chawner, B., & Oliver, G. (2013). A survey of New Zealand academic reference librarians: Current and future skills and competencies. *Australian Academic & Research Libraries*, *44*(1), 29-39.

Chen, Y., Jiang, S., & Zhou, L. A. (2020). Estimating returns to education in urban China: Evidence from a natural experiment in schooling reform. *Journal of Comparative Economics*, *48*(1), 218-233.

Chiswick, B. R. (1980). The earnings of white and coloured male immigrants in Britain. *Economica*, *47*(185), 81-87.

Christl, M., & Köppl–Turyna, M. (2020). Gender wage gap and the role of skills and tasks: evidence from the Austrian PIAAC data set. *Applied Economics*, *52*(2), 113-134.

Colecchia, A., & Schreyer, P. (2002). The contribution of information and communication technologies to economic growth in nine OECD countries. *OECD economic studies*, *2002*(1), 153-171.

Brown, C., Gay, P., & Policy Studies Institute. (1985). *Racial discrimination: 17 years after the Act*. Policy Studies Institute.

Coniglio, N. D., & Prota, F. (2008). Human capital accumulation and migration in a peripheral EU region: the case of Basilicata. *Papers in Regional Science*, *87*(1), 77-96.

Corcoran, J., Faggian, A., & McCann, P. (2010). Human capital in remote and rural Australia: The role of graduate migration. *Growth and Change*, *41*(2), 192-220.

Cortes, P., & Pan, J. (2018). Occupation and gender. *The Oxford handbook of women and the economy*, 425-452.

Cramer, L. (2021). Alternative strategies for closing the award gap between white and minority ethnic students. *eLife*, *10*, e58971. https://doi.org/10.7554/eLife.58971

Crawford, C., & Greaves, E. (2015). *Ethnic minorities substantially more likely to go to university than their White British peers*. Institute for Fiscal Studies.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Supplement to "estimating the technology of cognitive and noncognitive skill formation". *Econometrica*, *78*(3), 883-931.
Darrah, C. N. (2013). *Learning and work: An exploration in industrial ethnography*. Routledge.
De Noronha, N. (2022). *Race equality in Greater Manchester: An analysis of key issues*. Centre on the Dynamics of Ethnicity (CoDE), University of Manchester.
Dearden, L. (1999). The effects of families and ability on men's education and earnings in Britain. *Labour Economics*, *6*(4), 551-567.
Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593-1640.
Denzler, S., & Wolter, S. C. (2011). Too far to go? Does distance determine study choices?.
Department for Education. (2020). Destinations of students after 16 to 18 study. https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/after-education/destinations-of-students-after-key-stage-5-usually-aged-18-years/latest/#by-ethnicity
Dickerson, A., & Green, F. (2004). The growth and valuation of computing and other generic skills. *Oxford economic papers*, *56*(3), 371-406.
Dickerson, A., & Morris, D. (2019). The changing demand for skills in the UK. *Centre for Vocational Education Research, Research Paper*, *20*.
DiNardo, J. (2002). *Propensity score reweighting and changes in wage distributions*. University of Michigan. Mimeograph.
DiNardo, J. E., & Pischke, J. S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too?. *The Quarterly Journal of Economics*, *112*(1), 291-303.
Dolton, P., & Makepeace, G. (2004). Computer use and earnings in Britain. *The Economic Journal*, *114*(494), C117-C129.
Dolton, P., & Vignoles, A. (1999). *The Economic Case for Reforming A Levels* (No. dp0422). Centre for Economic Performance, LSE.
Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of education review, 19*(2), 179-198.
Donnelly, M., & Gamsu, S. (2018). *Home and away: Social, ethnic and spatial inequalities in student mobility*. The Sutton Trust.
Dotti, N. F., Fratesi, U., Lenzi, C., & Percoco, M. (2013). Local labour markets and the interregional mobility of Italian university students. *Spatial Economic Analysis*, *8*(4), 443-468.
Drinkwater, S., & Blackaby, D. (2004). *Migration and labour market differences: The case of Wales*. Available at SSRN: https://ssrn.com/abstract=585309
Drummond, I., Nixon, I., & Wiltshire, J. (1998). Personal transferable skills in higher education: The problems of implementing good practice. *Quality assurance in education*, *6*(1), 19-27.
Dunford, M., & Fielding, A. J. (1997). Greater London, the South-East region, and the wider Britain: Metropolitan polarization, uneven development and inter-regional migration. In *Regional policy and development* (274–293). *Regional Studies Association*.

Edin, P. A., Fredriksson, P., Nybom, M., & Öckert, B. (2022). The rising return to noncognitive skill. *American Economic Journal: Applied Economics*, *14*(2), 78-100.
Elias, P., & Purcell, K. (2004). Is mass higher education working? Evidence from the labour market experiences of recent graduates. *National Institute Economic Review, 190*, 60-74.
Elliott, R. J., & Lindley, J. K. (2008). Immigrant wage differentials, ethnicity and occupational segregation. Journal of the Royal Statistical Society Series A: *Statistics in Society, 171*(3), 645-671.
Entorf, H., & Kramarz, F. (1997). Does unmeasured ability explain the higher wages of new technology workers?. *European Economic Review*, *41*(8), 1489-1509.
Entorf, H., & Kramarz, F. (1998). The impact of new technologies on wages: lessons from matching panels on employees and on their firms. *Economics of innovation and new technology*, *5*(2-4), 165-198.
Evans, T. (2020, October 12). Ethnicity pay gaps: 2019. Office for National Statistics. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/articles/ethnicitypaygapsingreatbritain/2019
Faggian, A., & McCann, P. (2006). Human capital flows and regional knowledge assets: a simultaneous equation approach. *Oxford Economic Papers*, *58*(3), 475-500.
Faggian, A., Corcoran, J., & McCann, P. (2013). Modelling geographical graduate job search using circular statistics. *Papers in Regional Science*, *92*(2), 329-344.
Faggian, A., Mameli, F., & McCann, P. (2008). Employment growth in Italian local labour systems: Issues of model specification and sectoral aggregation. *Spatial economic analysis*, *3*(3), 343-360.
Faggian, A., McCann, P., & Sheppard, S. (2007). Human capital, higher education and graduate migration: an analysis of Scottish and Welsh students. *Urban Studies*, *44*(13), 2511-2528.
Felstead, A., Gallie, D., & Green, F. (2002). *Work skills in Britain 1986–2001*. DFES Publications.
Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness?. *The economic journal*, *114*(497), 641-659.
Firpo, S. P., Fortin, N. M., & Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics*, *6*(2), 28.
Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953-973.
Florida, R., Mellander, C., Stolarick, K., & Ross, A. (2012). Cities, skills and wages. *Journal of Economic Geography*, *12*(2), 355-377.
Form, W. (1987). On the degradation of skills. *Annual Review of Sociology*, *13*(1), 29-47.
Fortin, N., Lemieux, T., & Firpo, S. (2011). decomposition methods in economics. *In Handbook of labor economics*, *4*, 1-102. Elsevier.
Fulden, K. (2021). What Do Conditional and Unconditional Quantile Regression Models Tell Us Something Different About Wage Inequality in Turkey?. *Journal of Economy Culture and Society*, (64), 257-277.
Gallie, D., Felstead, A., & Green, F. (2004). Changing patterns of task discretion in

Britain. *Work, Employment and Society*, *18*(2), 243-266.
Gamsu, S., Donnelly, M., & Harris, R. (2019). The spatial dynamics of race in the transition to university: Diverse cities and White campuses in UK higher education. *Population, Space and Place, 25(*5), e2222.
Gimenez, R., De Oliveira, D. L., Manoel, E. D. J., Dantas, L., & Marques, I. (2012). Integrating fundamental movement skills in late childhood. *Perceptual and Motor Skills*, *114*(2), 563-583.
Gimpelson, V., Kapeliushnikov, R., & Lukiyanova, A. (2010). Stuck between surplus and shortage: demand for skills in Russian industry. *Labour, 24*(3), 311-332.
Glaeser, E. L., & Maré, D. C. (2001). Cities and skills. *Journal of labor economics*, *19*(2), 316-342.
Gonzalez, H. B., & Kuenzi, J. J. (2012). *Science, technology, engineering, and mathematics (STEM) education: A primer*. Congressional Research Service, Library of Congress.
Gopinath, G. (2020, April 14). The Great Lockdown: Worst economic downturn since the Great Depression. *IMF Blog*. https://www.imf.org/en/Blogs/Articles/2020/04/14/blog-weo-the-great-lockdown-worst-economic-downturn-since-the-great-depression
Gordon, I. (1991). *The impact of economic change on minorities and migrants in Western Europe*. Department of Geography, University of Reading.
Gordon, I. R., & Molho, I. (1998). A Multi-stream Analysis of the Changing Pattern of Interregional Migration in Great Britain, 1960-1991. *Regional Studies, 32*(4), 309-323.
Gordon, M. M. (1978). *Human nature, class, and ethnicity*. Oxford University Press.
Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *The Review of Economic Studies*, *74*(2), 477-506.
Great Britain. Department for Business, Energy and Industrial Strategy. (2017). *Industrial strategy: Building a Britain fit for the future*. HM Government.
Green, F. (1998). *The value of skills* (No. 9819). Department of Economics Discussion Paper.
Green, F., & McIntosh, S. (2007). Is there a genuine under-utilization of skills amongst the over-qualified?. *Applied economics*, *39*(4), 427-439.
Green, F., & Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, *62*(4), 740–763.
Green, F., McIntosh, S., & Vignoles, A. (2002). The utilization of education and skills: evidence from Britain. *The Manchester School*, *70*(6), 792-811.
Green, S. L., & Weinberg, R. S. (2001). Relationships among athletic identity, coping skills, social support, and the psychological impact of injury in recreational participants. *Journal of applied sport psychology*, *13*(1), 40-59.
Greenacre, M., Groenen, P. J., Hastie, T., d'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, *2*(1), 100.
Greenan, N., Kalugina, E., & Walkowiak, E. (2014). Has the quality of working life improved in the EU-15 between 1995 and 2005?. *Industrial and Corporate Change*, *23*(2), 399-428.
Greenan, N., Narcy, M., & Robin, S. (2013). *Changes Within Companies and Access of Older Workers to Further Training in the 1990's and 2000's*. HAL.

Groot, W., & Van Den Brink, H. M. (2002). Age and education differences in marriages and their effects on life satisfaction. *Journal of Happiness Studies, 3*(2), 153–165.

Guryan, J., & Charles, K. K. (2013). Taste‐based or statistical discrimination: the economics of discrimination returns to its roots. *The Economic Journal, 123*(572), 417-432.

Haisken-DeNew, J. P., & Schmidt, C. M. (1999). Money for nothing and your chips for free? The anatomy of the PC wage differential. *The Anatomy of the Pc Wage Differential (December 1999)*.

Handel, M. J. (2000). *Trends in direct measures of job skill requirements* (Working Paper No. 301). Levy Economics Institute of Bard College.

Handel, M. J. (2020). *Job skill requirements: Levels and trends* (MPRA Paper No. 100590). University Library of Munich, Germany. https://mpra.ub.uni-muenchen.de/100590/

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

Healy, G., Bradley, H., & Forson, C. (2011). Intersectional sensibilities in analysing inequality regimes in public sector organizations. *Gender, Work & Organization, 18*(5), 467-487.

Heath, A., & Cheung, S. Y. (2006). *Ethnic penalties in the labour market: Employers and discrimination* (Department for Work and Pensions Research Report No. 341). Corporate Document Services.

Heath, A., & Li, Y. (2008). Period, life-cycle, and generational effects on ethnic minority success in the British labour market. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *60*(48), 77-97. https://doi.org/10.1007/s11577-008-0061-7

Higher Education Statistics Agency. (2020). *UK domiciled undergraduate students of known ethnicity by subject area, ethnicity and academic year*. https://www.hesa.ac.uk/data-and-analysis/students/Table-24

Hoare, A., & Corver, M. (2010). The regional geography of new young graduate labour in the UK. *Regional Studies*, *44*(4), 477-494.

International Labour Office. (2010). *A skilled workforce for strong, sustainable and balanced growth: A G20 training strategy*. International Labour Office.

Jackson, D. (2016). Modelling graduate skill transfer from university to the workplace. *Journal of Education and Work*, *29*(2), 199-231.

Jackson, J. S., Hudson, D., Kershaw, K., Mezuk, B., Rafferty, J., & Tuttle, K. K. (2011). Discrimination, chronic stress, and mortality among black Americans: A life course framework. In J. E. Smith & D. R. Jones (Eds.), *International handbook of adult mortality*, 311-328. Springer.

Johnson, R. C. (2007). Wage and job dynamics after welfare reform: The importance of job skills. In *Aspects of Worker Well-Being* (pp. 231–298). Emerald Group Publishing Limited. https://doi.org/10.1016/S0147-9121(06)28006-9

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy, 87*(5, Part 1), 972-990.

Jung, J., & Chan, C. K. Y. (2017). Academics' perception on research versus teaching and their recognition. *The changing academic profession in Hong Kong*, 145-160.

Kain, J. F. (1968). Housing segregation, Negro employment, and metropolitan decentralization. *Quarterly Journal of Economics, 82*(2), 175–197.

Kamhöfer, D. A., & Schmitz, H. (2016). Reanalyzing zero returns to education in Germany. *Journal of Applied Econometrics, 31*(5), 912-919.

Kenny, L. W., Lee, L. F., Maddala, G. S., & Trost, R. P. (1979). Returns to college education: An investigation of self-selection bias based on the project talent data. *International Economic Review, 20*(3), 775–789.

Kidd, M., O'Leary, N. C., & Sloane, P. J. (2017). The impact of mobility on early career earnings: A quantile regression approach for UK graduates. *Economic Modelling, 62,* 92-102.

Kieffer, K. M. (1998). Orthogonal versus oblique factor rotation: A review of the literature regarding the pros and cons. In *Proceedings of the Annual Meeting of the Mid-South Educational Research Association*, 4-6.

Kiker, B. F., Santos, M. C., & De Oliveira, M. M. (1997). Overeducation and undereducation: Evidence for Portugal. *Economics of Education Review, 16*(2), 111-125.

Koc, E., & Koncz, A. (2009). Job outlook 2010. *Bethlehem, PA: National Association of Colleges and Employers*.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, *46*(1), 33–50.

Krueger, N. (1993). The impact of prior entrepreneurial exposure on perceptions of new venture feasibility and desirability. *Entrepreneurship theory and practice*, *18*(1), 5-21.

Kupets, O. (2016). Education-job mismatch in Ukraine: Too many people with tertiary education or too many jobs for low-skilled?. *Journal of comparative economics, 44*(1), 125-147.

Lang, K., & Lehmann, J. Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, *50*(4), 959-1006.

Larsen, J. K., & Wigand, R. T. (1987). Industry-university technology transfer in microelectronics. *Review of Policy Research, 6*(3), 584-595.

Leckey, J. F., & McGuigan, M. A. (1997). Right tracks—wrong rails: The development of generic skills in higher education. *Research in Higher Education*, *38*(3), 365–378.

Leslie, D., Drinkwater, S., & O’Leary, N. (1998). Unemployment and earnings among Britain’s ethnic minorities: some signs for optimism. *Journal of Ethnic and Migration Studies*, *24*(3), 489–506.

Leung, K. C., Leung, F. K., & Zuo, H. (2014). A study of the alignment of learning targets and assessment to generic skills in the new senior secondary mathematics curriculum in Hong Kong. *Studies in Educational Evaluation*, *43*, 115-132.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2), 139-191.

Li, Y., & Heath, A. (2020). Persisting disadvantages: a study of labour market dynamics of ethnic unemployment and earnings in the UK (2009–2015). *Journal of Ethnic and Migration Studies*, *46*(5), 857-878.

Lindley, J. (2002). Race or religion? The impact of religion on the employment and earnings of Britain's ethnic communities. *Journal of Ethnic and Migration Studies, 28*(3), 427-422.

Lindley, J. (2009). The over-education of UK immigrants and minority ethnic groups: Evidence from the Labour Force Survey. *Economics of Education Review, 28*(1), 80-89.

Longhi, S., & Brynin, M. (2017). *The ethnicity pay gap*. Equality and Human Rights Commission. Retrieved from https://www.equalityhumanrights.com

Longhi, S., Nicoletti, C., & Platt, L. (2012). Interpreting wage gaps of disabled men: The roles of productivity and of discrimination. *Southern Economic Journal, 78*(3), 931-953.

Machado-da-Silva, C. L., Fonseca, V. S. D., & Crubellate, J. M. (2005). Unlocking the institutionalization process: Insights for an institutionalizing approach. *BAR-Brazilian Administration Review, 2*(1), 1–20.

Mahuteau, S., Mavromaras, K., Sloane, P., & Wei, Z. (2014). *Horizontal and vertical educational mismatch and wages*. Paper presented at 25th Australian Labour Market Research.

Mameli, F., Faggian, A., & McCann, P. (2008). Employment growth in Italian local labour systems: Issues of model specification and sectoral aggregation. *Spatial economic analysis, 3(*3), 343-360.

Markaki, Y., & Longhi, S. (2013). What determines attitudes to immigration in European countries? An analysis at the regional level. *Migration Studies, 1*(3), 311-337.

Mason, G., O'Mahony, M., & Riley, R. (2018). What is Holding Back UK Productivity? Lessons from Decades of Measurement. *National Institute Economic Review, 246*(1), 24-35.

Mavromaras, K., & McGuinness, S. (2012). Overskilling dynamics and education pathways. *Economics of Education Review, 31*(5), 530-539.

Mavromaras, K., Sloane, P., & Wei, Z. (2012). The role of education pathways in the relationship between job mismatch, wages, and job satisfaction: A panel estimation approach. *Education Economics, 20*(3), 311-334.

McCann, P., & Sheppard, S. (2001). Public investment and regional labour markets: The role of UK higher education. In M. Madden (Ed.), *Public investment and regional economic development: Essays in honour of Moss Madden* (135-153). Edward Elgar Publishing.

McGovern, P. (2007). Immigration, labour markets and employment relations: Problems and prospects. *British Journal of Industrial Relations, 45*(2), 217-235.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of economic surveys, 20*(3), 387-418.

Mellander, C., Florida, R., & Stolarick, K. (2011). Here to stay—the effects of community satisfaction on the decision to stay. *Spatial Economic Analysis*, *6*(1), 5-24.

Melly, B., & Wüthrich, K. (2017). Local quantile treatment effects. In A. M. Ellison & J. M. Anderson (Eds.), *Handbook of quantile regression* (145-164). Springer.

Metcalf, D. (2013). *Immigration and the British labour market: The role of the Migration Advisory Committee*. Migration Advisory Committee. London School of Economics.

Metcalf, H. (2009). *Pay gaps across the equality strands: A review*. Equality and Human Rights Commission. Retrieved from https://www.equalityhumanrights.com

Mincer, J. (1975). Education, experience, and the distribution of earnings and employment: An overview. In J. S. Coleman (Ed.), *Education, income, and human behavior* (71-94). Columbia University Press.

Mincer, J. (1989). *Human capital responses to technological change in the labor market* (NBER Working Paper No. w3207). National Bureau of Economic Research. Available at https://ssrn.com/abstract=226714

Mincer, J. A. (1974). The human capital earnings function. In *Schooling, experience, and earnings* (83–96). National Bureau of Economic Research.

Mirza, H. S. (2006). 'Race', gender, and educational desire. *Race Ethnicity and Education, 9*(2), 137–158.

Mosca, I., & Wright, R. E. (2010). National and international graduate migration flows. *Population Trends, 141*, 36–53.

Murphy, P., Blackaby, D., O'Leary, N., & Staneva, A. (2020). Understanding what has been happening to the public-sector pay premium in Great Britain: A distributional approach based on the Labour Force Survey. *British Journal of Industrial Relations, 58*(2), 273–300.

Musset, P., & Field, S. (2013). *A skills beyond school review of England*. Paris: OECD Publishing.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, *13*(4), 653-677.

Macintosh, N. W. G., & Smith, D. J. (1974). *The extent of racial discrimination*. London: Political and Economic Planning

Niven, J., Faggian, A., & Ruwanpura, K. N. (2013). Exploring "underachievement" among highly educated young British-Bangladeshi women. *Feminist Economics*, *19*(1), 111–136.

Nordin, M., Persson, I., & Rooth, D. O. (2010). Education–occupation mismatch: Is there an income penalty?. *Economics of education review, 29*(6), 1047-1059.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693–709.

Okafor, E. I., Amunnadi Chukwudi, A., Otiji, G. O., Onyia, B. N., & Chikelu, J. U. (2023). Feminism in Chimamanda's americana: guidance/counselling and education management on the intersection between feminism and African tradition. *Igwebuike: African Journal of Arts and Humanities, 9*(1), 53-66.

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

O'Leary, N. C., & Sloane, P. J. (2008). Rates of return to degrees across British regions. *Regional Studies*, *42*(2), 199-213.

O'Leary, N. C., Murphy, P. D., & Blackaby, D. H. (2004). Quantile regression estimates of the union wage effect for Great Britain. *The Manchester School, 72*(4), 497-514.

Oulton, N. (2001). Why do foreign-owned firms in the UK have higher labour productivity? In N. Pain (Ed.), *Inward investment, technological change, and growth: The impact of multinational corporations on the UK economy* (pp. 122–161). Palgrave Macmillan.

Potrac, P., & Jones, R. (2009). Power, conflict, and cooperation: Toward a micropolitics of coaching. *Quest*, *61*(2), 223-236.
Quintini, G. (2014). *Skills at work: How skills and their use matter in the labour market* (OECD Social, Employment and Migration Working Papers No. 158). OECD Publishing. https://doi.org/10.1787/5jz44fdfjm7j-en
Ramos, J. L. S., Dolipas, B. B., & Villamor, B. B. (2013). Higher order thinking skills and academic performance in physics of college students: A regression analysis. *International Journal of Innovative Interdisciplinary Research*, *4*(1), 48-60.
Rios-Avila, F., & de New, J. P. (2022). *Marginal unit interpretation of unconditional quantile regression and recentered influence functions using centred regression* (Melbourne Institute Working Paper No. 14/22). Melbourne Institute: Applied Economic & Social Research, The University of Melbourne.
Riveros, L. A., Bouton, L., & Mundial, B. (1991). *Efficiency wage theory, labor markets, and adjustment* (Vol. 731). Population and Human Resources Department, The World Bank.
Roberts, N., & Bolton, P. (2020). *Educational outcomes of Black pupils and students* (Briefing Paper No. 9023). House of Commons Library.
Robst, J. (2006). Estimation of a Hedonic Pricing Model for Medigap Insurance. *Health services research*, *41*(6), 2097-2113.
Robst, J. (2007). Education and job match: The relatedness of college major and work. *Economics of Education Review, 26*(4), 397-407.
Rosenberg, S., Heimler, R., & Morote, E. S. (2012). Basic employability skills: A triangular design approach. *Education+ Training*, *54*(1), 7-20.
Rothwell, N., Bollman, R. D., Tremblay, J., & Marshall, J. (2002). *Migration to and from rural and small town Canada*. Statistics Canada, Agriculture Division.
Rotundo, M., & Sackett, P. R. (2004). Specific versus general skills and abilities: A job level examination of relationships with wage. *Journal of Occupational and Organizational Psychology*, *77*(2), 127-148.
Rubb, S. (2003). Overeducation in the labor market: a comment and re-analysis of a meta-analysis. *Economics of Education review, 22*(6), 621-629.
Rudakov, V., Figueiredo, H., Teixeira, P., & Roshchin, S. (2021). The impact of horizontal job-education mismatches on the earnings of recent university graduates in Russia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3726368
Rumberger, R. W. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources, 22*(3), 24–50.
Salas Velasco, M. (2014). Do higher education institutions make a difference in competence development? A model of competence production at university. *Higher Education, 68*(4), 503–523.
Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature, 31*(2), 831–880.
Savic, M., Vecchi, M., & Lewis, A. (2019). *Overeducation and hourly wages in the UK labour market: 2006 to 2017*. Office for National Statistics.
Schwab, J. A. (2002). *Multinomial logistic regression: Basic relationships and complete problems*. Austin, TX: University of Texas.

Shiner, M., & Modood, T. (2002). Help or hindrance? Higher education and the route to ethnic equality. *British journal of sociology of education*, *23*(2), 209-232.

Simmie, J., & Strambach, S. (2006). The contribution of KIBS to innovation in cities: an evolutionary and institutional perspective. *Journal of knowledge management*, *10*(5), 26-40.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy, 70*(5, Part 2), 80–93. https://doi.org/10.1086/258726

Sloane, P. J. (2003). Much ado about nothing? What does the overeducation literature really tell us? In *Overeducation in Europe* (pp. 11–46). Economics Archive B.

Sloane, P., O'Leary, N., Murphy, P., & Blackaby, D. (2003). *Returns to education: A survey of findings*. Report for Welsh Assembly Government.

Spence, M. (1978). Job market signaling. In *Uncertainty in economics* (281-306). Academic Press.

Spenner, K. I. (1990). Skill: meanings, methods, and measures. *Work and occupations*, *17*(4), 399-421.

Star, C., & Hammer, S. (2008). Teaching generic skills: Eroding the higher purpose of universities, or an opportunity for renewal?. *Oxford Review of Education*, *34*(2), 237-251.

Stevens, M. (1994). A theoretical model of on-the-job training with imperfect competition. *Oxford Economic Papers*, 46(4), 537–562.

Sung, Y., Turner, S. L., & Kaewchinda, M. (2013). Career development skills, outcomes, and hope among college students. *Journal of career Development*, *40*(2), 127-145.

Sweetman, R., Hovdhaugen, E., & Karlsen, H. (2014). Learning outcomes across disciplinary divides and contrasting national higher education traditions. *Tertiary Education and Management*, *20*, 179-192.

Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA* (Vol. 724). Belmont, CA: Thomson/Brooks/Cole.

Takhar, O. K. (2016). *Sikh identity: An exploration of groups among Sikhs*. Routledge.

Todaro, M. P. (1969). A model of labor migration and urban unemployment in less developed countries. *The American Economic Review*, 59(1), 138-148.

Trostel, P., Walker, I., & Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour economics*, *9*(1), 1-16.

Tsang, M. C., Rumberger, R. W., & Levin, H. M. (1991). The impact of surplus schooling on worker productivity. *Industrial Relations: A Journal of Economy and Society, 30*(2), 209–228.

Vecchi, M., Robinson, C., Savic, M., & Romiti, M. (2023). *Vertical and horizontal mismatch in the UK: Are graduates' skills a good fit for their jobs?* National Institute of Economic and Social Research (NIESR) Discussion Papers, 548.

Velasco, V., Gragnano, A., Gruppo Regionale HBSC Lombardia 2018, & Vecchio, L. P. (2021). Health literacy levels among Italian students: Monitoring and promotion at school. *International Journal of Environmental Research and Public Health, 18*(19), 9943.

Venhorst, V., Van Dijk, J., & Van Wissen, L. (2011). An analysis of trends in spatial mobility of Dutch graduates. *Spatial Economic Analysis*, *6*(1), 57-82.

Walby, S., & Armstrong, J. (2010). Measuring equalities: Data and indicators in Britain. *International Journal of Social Research Methodology, 13*(3), 243–256.
Walker, I., & Zhu, Y. (2007). *The labour market effects of qualifications*. Futureskills Scotland Research Series.
Walker, I., & Zhu, Y. (2008). The college wage premium and the expansion of higher education in the UK. *The Scandinavian Journal of Economics, 110*(4), 695-709.
Walker, I., & Zhu, Y. (2011). Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review, 30*(6), 1177-1186.
Winters, J. V. (2020). In-state college enrollment and later life location decisions. *Journal of Human Resources, 55*(4), 1400-1426.
Wood, M., Hales, J., Purdon, S., Sejersen, T., & Hayllar, O. (2009). *A test for racial discrimination in recruitment practice in British cities* (Research Report No. 607). Department for Work and Pensions.
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
Zutshi, S. (2008). Differing immigrant wages: Causes for differences in immigrants' standards of living. *Empirical Economic Bulletin: An Undergraduate Journal, 1*(1), 2.
Zwysen, W., & Longhi, S. (2016). *Labour market disadvantage of ethnic minority British graduates: University choice, parental background or neighbourhood?* (Working Paper No. 2016-02). ISER Working Paper Series.
Zwysen, W., Di Stasio, V., & Heath, A. (2021). Ethnic penalties and hiring discrimination: Comparing results from observational studies with field experiments in the UK. *Sociology, 55*(2), 263–282.