

**ESSAYS ON THE LABOUR MARKET OUTCOMES OF  
IMMIGRANTS IN THE UK**

ZOVANGA LOUIS KONE

Thesis submitted to the University of Nottingham for the degree of  
Doctor of Philosophy

June, 2016

Dedicated to N’Nan Yah Kone

who I am very fortunate to have as a grandmother

## **Abstract**

The present thesis comprises three essays on the labour market outcomes of immigrants in the United Kingdom (UK).

Chapter 1 introduces the thesis, outlines its contribution and provides an overview of each of the three essays.

Chapter 2 evaluates the role of immigration policy on the occupational outcomes of immigrants to the UK. We use a quasi-natural experiment to tease out the role of immigration policy on the occupational outcomes of immigrants from other factors that affect these outcomes. Looking at a sample of immigrants who entered the UK shortly before and after 2004, the findings suggest that the change in the UK's 2004 immigration policy only led to a slight increase, although not statistically robust, in the odds of observing A8 immigrants who entered the UK after 2004 in elementary occupations relative to professional occupations. We unearth evidence that A8 immigrants with longer durations of stay in the country have better occupational outcomes. Also, the occupational attainment gap between A8 immigrants and UK-born individuals reduces drastically once one controls for income levels in the source country of the immigrant.

In Chapter 3, we examine the implications of networks of social contacts for the occupational outcomes of immigrants with different lengths of stay in the host country. It is commonly assumed that immigrants principally rely on co-ethnics to find employment. Using a direct measure on whether employment was obtained by referral or other means, we find that while this common assumption remains true for immigrants who recently arrived in the UK, co-ethnics no longer appear to be the main source of referrals for more established immigrants. Thus it seems that the social network of contacts of an immigrant potentially extends to include individuals from other countries as their stay in the host country

lengthens. This implies that immigrants may undergo a process of “social assimilation” by broadening their network of social contacts. If this diversification improves the “quality” of the network, it could lead to better labour market outcomes.

In Chapter 4, we look at the labour market outcomes of different groups of immigrants and the children born to them in the UK, in comparison to natives. Most previous studies assume that all UK-born of non-White ethnicity are children of immigrants because the data commonly used do not identify the place of birth of the respondent’s parents. In light of the history of immigration of some ethnic groups in the UK, however, such an assumption may lead to classification errors in the data, which could have severe consequences for implications of intergenerational mobility in labour market outcomes. Our analysis shows that even an apparently negligible amount of classification errors in the data can cause high level of uncertainty in the estimates of the parameters of interest. Additionally, for the first time in this literature we use a dataset that contains information on parental country of birth to examine labour market outcomes across all major ethnic groups in the UK.

Chapter 5 concludes the thesis, and future research topics/directions are discussed.

## Acknowledgments

I have benefitted from the kind support, encouragement and advice of many throughout my doctoral studies. I cannot name all of them for it would be a lengthy list, but I remain grateful to each of them.

First and foremost, I would like to thank my supervisors, Professor Sourafel Girma and Professor Richard Upward, for their unmeasured support and advice. Their experiences and insights from conducting research over the years were made freely available to me, and I have benefitted greatly. I also am indebted to Professor Richard Disney for being a valuable source of advice.

I am thankful to Professor Oliver Morrissey for his comments on my first draft of Chapter 4. Many words of gratitude go to the Staff of the School of Economics – both academic and support. The thesis benefitted from comments received at the School of Economics PhD Annual Conferences. I thank all the participants. I also thank friends and PhD colleagues (especially Frank, Kian, Mo, Nihar, Priscilla, Sameen, Tim and Veronica) for all the discussions – some long, some short – which, directly or indirectly, added value to the thesis.

I am grateful to my parents and siblings who never ceased to care for my wellbeing during my studies – well, they still do. I am blessed to have them.

I benefitted from the generous funding of the Economics and Social Research Council (ESRC) for my doctoral studies. Their support is highly appreciated.

# Table of Contents

|                                                                                                                                                    |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Abstract .....                                                                                                                                     | ii  |
| Acknowledgments.....                                                                                                                               | iv  |
| List of tables.....                                                                                                                                | vii |
| List of figures.....                                                                                                                               | ix  |
| Chapter 1: Introduction .....                                                                                                                      | 1   |
| 1.1 Immigrants in the UK in recent years .....                                                                                                     | 2   |
| 1.2 The labour market outcomes of immigrants.....                                                                                                  | 4   |
| 1.3 Outline of the thesis.....                                                                                                                     | 11  |
| Chapter 2: The role of immigration policy on the occupational outcomes of immigrants .....                                                         | 15  |
| 2.1 Introduction.....                                                                                                                              | 16  |
| 2.2 Institutional context .....                                                                                                                    | 18  |
| 2.3 Related literature .....                                                                                                                       | 23  |
| 2.4 Ex-ante expectation of the implications of the relaxation of visa requirements. ....                                                           | 27  |
| 2.5 The data.....                                                                                                                                  | 32  |
| 2.6 Methodology .....                                                                                                                              | 37  |
| 2.6.1 Analysis of occupational outcomes.....                                                                                                       | 37  |
| 2.6.2 Disentangling the impact of immigration policy from the impact of birthplace on the occupational outcomes of immigrants.....                 | 38  |
| 2.6.3 Challenges of identifying the impact of immigration policy on the occupational outcome of immigrants based on independent cross section..... | 40  |
| 2.6.4 The impact of policy on occupational outcomes.....                                                                                           | 42  |
| 2.7 Empirical analysis and results .....                                                                                                           | 43  |
| 2.7.1 Plausibility of the underlying of assumptions in section 2.6.3.....                                                                          | 47  |
| 2.7.2 Immigration policy and the occupational outcomes of immigrants .....                                                                         | 53  |
| 2.7.3 Further robustness check: sensitivity to identity of the control group .....                                                                 | 62  |
| 2.8 Discussion.....                                                                                                                                | 65  |
| 2.9 Conclusion.....                                                                                                                                | 71  |
| 2.10 Appendix.....                                                                                                                                 | 73  |
| 2.10.1 Requirements for registering on the Worker Registration Scheme (WRS). ....                                                                  | 73  |
| Chapter 3: Job-finding networks and migrants' occupational outcomes.....                                                                           | 76  |
| 3.1 Introduction.....                                                                                                                              | 77  |
| 3.2 Literature Review.....                                                                                                                         | 81  |

|                                                                                                                       |                                                                                                                            |     |
|-----------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|-----|
| 3.2.1                                                                                                                 | Defining and measuring networks .....                                                                                      | 82  |
| 3.2.2                                                                                                                 | Identifying the network effect.....                                                                                        | 85  |
| 3.3                                                                                                                   | The data.....                                                                                                              | 87  |
| 3.4                                                                                                                   | Job networks and occupational outcomes.....                                                                                | 92  |
| 3.5                                                                                                                   | Methodology .....                                                                                                          | 96  |
| 3.5.1                                                                                                                 | The relationship between networks of social contacts and occupational outcomes .....                                       | 97  |
| 3.6                                                                                                                   | Econometric analysis .....                                                                                                 | 101 |
| 3.6.1                                                                                                                 | The size of networks in an occupation and the likelihood of finding employment in a given occupation through referral..... | 102 |
| 3.6.2                                                                                                                 | Replicating Patel and Vella (2013).....                                                                                    | 112 |
| 3.6.3                                                                                                                 | Extending Patel and Vella's approach .....                                                                                 | 116 |
| 3.7                                                                                                                   | Conclusion.....                                                                                                            | 121 |
| 3.8                                                                                                                   | Appendix.....                                                                                                              | 123 |
| Chapter 4: The impact of misclassification error on the measurement of the labour market performance of migrants..... |                                                                                                                            | 124 |
| 4.1                                                                                                                   | Introduction.....                                                                                                          | 125 |
| 4.2                                                                                                                   | Literature Survey.....                                                                                                     | 127 |
| 4.3                                                                                                                   | The data.....                                                                                                              | 131 |
| 4.3.1                                                                                                                 | Sample characteristics.....                                                                                                | 137 |
| 4.3.2                                                                                                                 | Source of misclassification error.....                                                                                     | 138 |
| 4.4                                                                                                                   | Methodology .....                                                                                                          | 140 |
| 4.4.1                                                                                                                 | Implications of misclassification.....                                                                                     | 140 |
| 4.5                                                                                                                   | Empirical Analysis .....                                                                                                   | 142 |
| 4.5.1                                                                                                                 | Estimation .....                                                                                                           | 142 |
| 4.6                                                                                                                   | Results .....                                                                                                              | 145 |
| 4.6.1                                                                                                                 | Education .....                                                                                                            | 145 |
| 4.6.2                                                                                                                 | Employment .....                                                                                                           | 147 |
| 4.6.3                                                                                                                 | Earnings.....                                                                                                              | 149 |
| 4.6.4                                                                                                                 | Allowing for 1 per cent of misclassification error in the LFS data used....                                                | 156 |
| 4.7                                                                                                                   | Conclusion.....                                                                                                            | 163 |
| 4.8                                                                                                                   | Appendix.....                                                                                                              | 165 |
| Chapter 5: Conclusion.....                                                                                            |                                                                                                                            | 171 |
| Bibliography.....                                                                                                     |                                                                                                                            | 176 |

## List of tables

|                                                                                                                                                                                                        |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 6.1: Mean age and mean age left continuous full-time education over the last two decades.....                                                                                                    | 4  |
| Table 6.2: Occupational distributions of immigrants and natives (in %), LFS (2013-2014) .....                                                                                                          | 6  |
| Table 6.3: Occupational distributions by group of immigrants (in %), LFS (2013-2014)....                                                                                                               | 7  |
| Table 2.1: Passenger admitted by purpose of journey, 1998 to 2008 (excludes EU nationals) .....                                                                                                        | 21 |
| Table 2.2: Classification of occupations based on skill requirements for task performed as per the ONS.....                                                                                            | 36 |
| Table 2.3: Occupational distributions by time of arrival and place of origin – LFS (2001-2009).....                                                                                                    | 36 |
| Table 2.4: Median age at arrival in the UK and educational attainment levels of immigrants, by year of arrival.....                                                                                    | 37 |
| Table 2.5: Summary statistic of the sample: Immigrants who entered the UK between 2001 and 2006, LFS (2001-2009).....                                                                                  | 46 |
| Table 2.6: Difference in means of characteristics before and after 2004, by immigrant group. Only immigrants who entered the UK between 2001 and 2006 .....                                            | 49 |
| Table 2.7: Test of differences in changes across groups reported in Table 2.6 .....                                                                                                                    | 49 |
| Table 2.8: Odds ratios estimated from a multinomial logit with professional occupations as base occupation; control group is non-EU immigrants.....                                                    | 58 |
| Table 2.9: Odds ratios estimated from a multinomial logit with professional occupations as base occupation. Sample excludes 2004 arrivals; control group is non-EU immigrants .....                    | 59 |
| Table 2.10: Odd ratios estimated from an ordered logit; control group is non-EU immigrants.....                                                                                                        | 62 |
| Table 2.11: Average marginal effects from LPM with professional occupations as base occupation; control group is non-EU immigrants .....                                                               | 62 |
| Table 2.12: Odds ratios estimated from a multinomial logit with professional occupations as base occupation; control group is immigrants from EU14 countries .....                                     | 64 |
| Table 2.13: Odds ratios estimated from a multinomial logit with professional occupations as base occupation. Roles of duration of stay in the UK and income levels in the migrant's home country ..... | 68 |
| Table 2.14: Difference in means of characteristics before and after 2004, by immigrant group. Only immigrants who entered the UK between 2001 and 2006. Sample excludes 2004 arrivals .....            | 74 |
| Table 2.15: Test of differences in changes in characteristics across groups reported in Table 2.14.....                                                                                                | 74 |

|                                                                                                                                                                                                                                                                                                         |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 2.16: Difference in means of characteristics before and after 2004, by immigrant group. Stock of immigrants in the UK between 2001 and 2006.....                                                                                                                                                  | 74  |
| Table 2.17: Test of differences in changes in characteristics across groups reported in Table 2.16.....                                                                                                                                                                                                 | 74  |
| Table 2.18: Placebo effect: looking at immigrants who entered the UK between 1998 and 2003; control group is non-EU immigrants .....                                                                                                                                                                    | 75  |
| Table 3.1: How job was obtained (in %) – new employees, LFS 1998-2009.....                                                                                                                                                                                                                              | 90  |
| Table 3.2: How job was obtained, by occupation (in %) – immigrants only, LFS 1998-2009.....                                                                                                                                                                                                             | 92  |
| Table 3.3: Template of the data before estimation of the Asclogit and the LPM.....                                                                                                                                                                                                                      | 100 |
| Table 3.4: Summary statistic of sample – employees with less than a year with employer .....                                                                                                                                                                                                            | 101 |
| Table 3.5: Likelihood of finding employment in an occupation through referral and the size of the network in the occupation: semi-elasticities from equations (3.6.1) and (3.6.2), reported across new employees with different durations of stay in the UK.....                                        | 109 |
| Table 3.6: Likelihood of finding employment in an occupation through referral and the size of the network in the occupation: semi-elasticities from equation (3.6.2), reported across new employees with different durations of stay in the UK. Restricted samples .....                                | 111 |
| Table 3.7: Likelihood of finding employment in an occupation through referral and the size of the network in the occupation: semi-elasticities from equation (3.6.2), reported across new employees with different durations of stay in the UK. Sample excludes those with a network size of zero ..... | 112 |
| Table 3.8: Robustness check with standard errors clustered by (1) country of birth; (2) country of birth and region of work. Average marginal effects from equation (3.6.2) using LPM, reported across new employees with different durations of stay in the UK. ....                                   | 112 |
| Table 3.9: Likelihood that a new employee chose the most popular occupation amongst established co-ethnics – average marginal effects from equations (3.6.3) and (3.6.4), reported across new employees with different durations of stay in the UK.....                                                 | 116 |
| Table 3.10: Likelihood of choosing a given occupation – marginal effects computed at the mean (MEM).....                                                                                                                                                                                                | 119 |
| Table 3.11: Probability of transition in employment status .....                                                                                                                                                                                                                                        | 123 |
| Table 3.12: Share of recent employees who found employment by “hearing from someone who worked there” (LFS, 2005 onward) .....                                                                                                                                                                          | 123 |
| Table 3.13: Likelihood of choosing a given occupation – average marginal effect (AME) of the migrant's network of social contacts from equation (3.6.5) .....                                                                                                                                           | 123 |
| Table 4.1: Summary statistics of sample, LFS (1993 to 2007) .....                                                                                                                                                                                                                                       | 134 |
| Table 4.2: Summary statistics of sample, GHS (1993 to 2006) .....                                                                                                                                                                                                                                       | 135 |
| Table 4.3: Breakdown of nativity status using the GHS.....                                                                                                                                                                                                                                              | 136 |
| Table 4.4: Source of misclassification.....                                                                                                                                                                                                                                                             | 136 |

|                                                                                                                                   |     |
|-----------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 4.5: Age left continuous full-time education – base group: natives .....                                                    | 152 |
| Table 4.6: Employment probabilities – base group: natives .....                                                                   | 153 |
| Table 4.7: Earnings (controlling for education) – base group: natives.....                                                        | 154 |
| Table 4.8: Earnings (not controlling for education) – base group: natives .....                                                   | 155 |
| Table 4.9: Relative Employment Probabilities with 1% misclassification error in the LFS .....                                     | 159 |
| Table 4.10: Relative Earnings (controlling for education) with 1% misclassification error in the LFS.....                         | 159 |
| Table 4.11: Comparing sample summary to those reported in ADGM .....                                                              | 165 |
| Table 4.12: Comparison to ADGM. Age left full-time education; base group: natives, LFS (93 to 07).....                            | 166 |
| Table 4.13: Comparison to ADGM. Employment probability; base group: natives, LFS (93 to 07) .....                                 | 167 |
| Table 4.14: Comparison to ADGM. Log net hourly earnings (controlling for education); base group: natives, LFS (93 to 07).....     | 168 |
| Table 4.15: Comparison to ADGM. Log net hourly earnings (not controlling for education); base group: natives, LFS (93 to 07)..... | 169 |
| Table 4.16: Characteristics of the 2 <sup>nd</sup> and 2.5 <sup>th</sup> generations, mean values .....                           | 170 |

## List of figures

|                                                                                                                                                                                                                |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 6.1: Share of immigrants in the UK over the last two decades .....                                                                                                                                      | 2   |
| Figure 6.2: Trend in age distributions of immigrants and natives.....                                                                                                                                          | 4   |
| Figure 6.3: Employment probability over time.....                                                                                                                                                              | 6   |
| Figure 2.1: Occupational distribution of A8 immigrants by duration of stay in the UK.....                                                                                                                      | 46  |
| Figure 2.2: Trend of occupational distributions over time, by group of immigrants: A8 and non-EU immigrants (vertical lines indicate 95% CI range).....                                                        | 51  |
| Figure 2.3: Trend of occupational distributions over time, by group of immigrants: A8 and non-EU immigrants (vertical lines indicate 95% CI range).....                                                        | 52  |
| Figure 3.1: share of new jobs found through networks of social contacts and share of new jobs in most popular occupations.....                                                                                 | 91  |
| Figure 4.1: Range of uncertainty arising from allowing 1% misclassification error introduced in the LFS sample.....                                                                                            | 160 |
| Figure 4.2: Relative employment probabilities – densities of point estimates with 1% misclassification error introduced in the LFS sample.....                                                                 | 161 |
| Figure 4.3: Range of uncertainty when comparing relative employment probabilities of 1 <sup>st</sup> and 2 <sup>nd</sup> generation immigrants, allowing for 1% misclassification error in LFS data used ..... | 162 |

## Chapter 1: Introduction

Globalisation has enabled, in addition to the increased variety in consumer choices, ease of movement of people across nations. In 2013, the stock of global immigrants stood at over 231 million compared with less than 80 million in 1960. The United Kingdom (UK) remains one of the most popular destination countries in the developed world for immigrants. With a birthplace diversity of migrants at 96, compared with 92 for the USA<sup>1</sup>, the UK has the most diverse pool of immigrants. The country has a distinct long history of being a host to immigrants:

“... There were French Jews in London, Lincoln, York and Norwich in the twelfth century; the Elizabeth age brought Italian musicians, German businessmen and even Africans, pressed by the first piratical stirring of the slave trade. Protestants from Low Countries, seeking religious tolerance, brought new trades and kick-started the Industrial Revolution in the same period. Huguenot refugees arrived from France *en masse* in the seventeenth century, creating a vibrant commercial atmosphere, as did the Greek Christians fleeing the Turks (a church was built in Soho for the Archbishop of Samos in 1677).

... Victorian Britain hummed with human traffic from all over Europe: scientists, engineers, teachers, labourers, maids, sailors and (following the failed revolutions of 1848) political dissidents and intellectuals...

... Immigration has given us many institutions that seem built into our heritage: Rothschilds, Warburgs, Barings, Reuters and Cazenove in the City; Marks & Spencer and Dolland & Aitchison in the high street; Trust House Forte and Tesco on ring roads... Then there are the literary figures who would not be here if they or their parents had not fixed their sights on Britain – T. S. Eliot, Joseph Conrad, Harold Pinter, Salman Rushdie, V. S. Naipaul, Germaine Greer, Doris Lessing and many others. The most prominent modern historian of these islands, Simon Schama, is himself the son of an immigrant. Immigration has given us shady tycoons (Robert Maxwell), respected generals (Peter de la Billière), noted intellectuals (George Steiner), amazing musicians (Alfred Brendel) and brilliant athletes (Linford Christie). ” [Robert Winder (2013), “Bloody Foreigners: The Story of Immigration to Britain”, p3, p8 & p9]

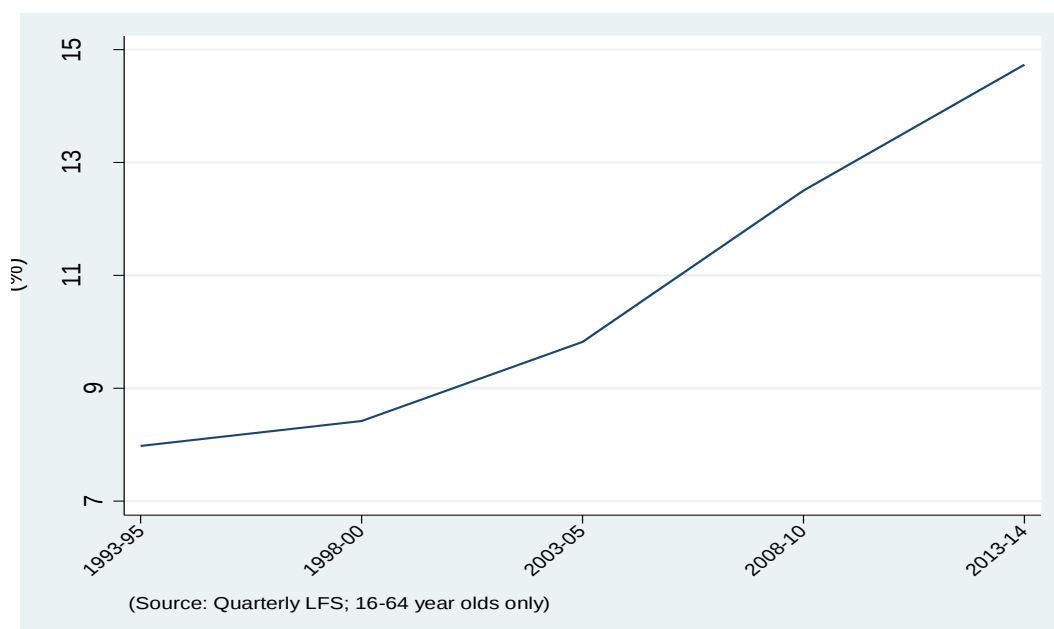
---

<sup>1</sup> See Alesina et al (2013), NBER Working paper 18699. The measure is based on an Herfindahl index of immigrant population group shares.

## 1.1 Immigrants in the UK in recent years

Figure 1.1 tells us that the UK is host to many foreign-born individuals – and to the children born to them. Over the past two decades, the share of foreign-born individuals of working age (i.e. 16 to 64) in the UK has steadily been increasing from about 8 per cent around the mid 1990's to almost 15 per cent in 2014. The diversity in the origin countries of immigrants in the UK several centuries ago, as highlighted by Robert Winder, still holds today. By 2014, the immigrant population in the UK was made up of individuals from 193 origin countries and territories. Europe remains the most popular region of origin of immigrants, accounting for over 37 per cent of foreign-born individuals. Those born in the Indian sub-continent (i.e. Bangladesh, India and Pakistan) account for close to 20 per cent, and the rest of Asia accounts for about 12 per cent. Sub-Saharan Africa is the birth place of over 16 per cent of immigrants. Over 3 per cent of immigrants were born in the Caribbean, and a little less than 3 per cent were born in the USA and Canada alone.

**Figure 1.1: Share of immigrants in the UK over the last two decades**



Immigrants tend to be younger than their native counterparts: the mean age of immigrants was 39 in 2014 in comparison to 41 amongst natives. Table 1.1 presents means of age for immigrants and natives over the last two decades.

Except for around the mid 1990's, we see from the upper left of the table that immigrants have consistently been younger than natives. This observation is illustrated in Figure 1.2 where we show how the age distributions of both immigrants and natives have changed over time. While the share of individuals between the ages of 25 and 44 has been declining amongst the latter, the corresponding share has been rising amongst immigrants. We also see that the share of those between the ages of 45 and 64 has been declining amongst immigrants; amongst natives, the corresponding share has seen a consistent increase. Such observations suggest that foreign-born individuals entering the UK are more likely to be in the prime of their working age.

Table 1.1 includes another measure of human capital, i.e. educational attainment, separately for immigrants and natives. We use age of completion of continuous full-time education as a measure of educational attainment because a large proportion of immigrants have their qualifications reported as “other” in the LFS, our data source – see Drinkwater et al (2009) and Algan et al (2010) who make similar remarks. We see that immigrants, on average, leave continuous full-time education at an older age than their native counterparts, as reported in the upper right of the table.

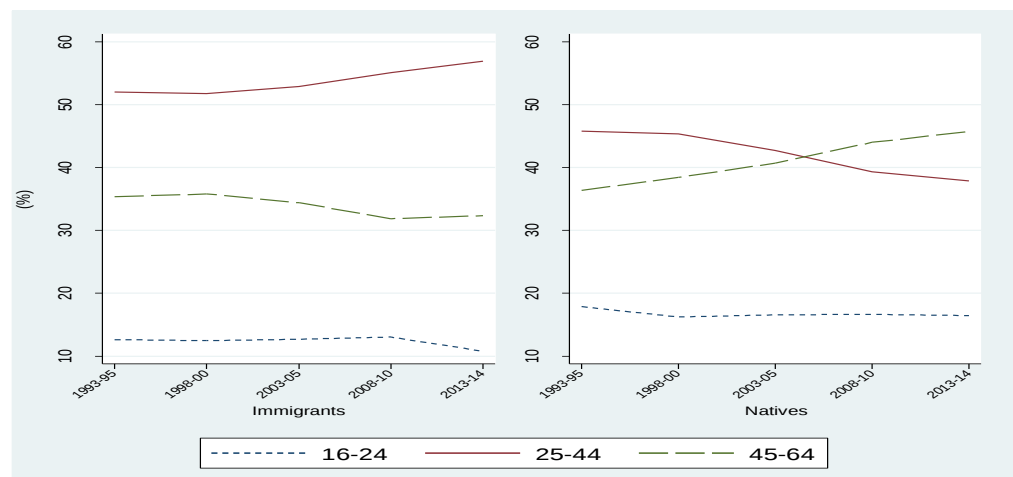
Means of age and educational attainment levels for different groups of immigrants are reported in the bottom half of Table 1.1. Immigrants are aggregated into three groups according to their place of birth: those from countries that were member states of the European Union (EU) before 2004 (i.e. EU14), those from Eastern and Central European countries that joined the EU in 2004 (i.e. A8) and those from other parts of the world (i.e. non-EU countries). We observe some heterogeneity in the reported characteristics across groups of immigrants. For example, while the mean age for A8 immigrants has been lower than those of the other two groups over the last 15 years, EU14 immigrants have had the highest mean age during the same period. We also see differences in educational attainment levels, although the extent of these are smaller in comparison to age. When one turns to within group comparison over time, it appears that there was a sudden drop in the mean age of A8 immigrants from the mid-2000's, the time A8 countries became member states of the EU and their nationals obtained virtually free access to the UK's labour market.

**Table 1.1: Mean age and mean age left continuous full-time education over the last two decades**

|         | Age                                   |                 | Age left continuous full-time education |                  |                 |                 |
|---------|---------------------------------------|-----------------|-----------------------------------------|------------------|-----------------|-----------------|
|         | Immigrants                            | Natives         | Immigrants                              | Natives          |                 |                 |
| 1993-95 | 39.66<br>(0.04)                       | 38.75<br>(0.01) | 17.93<br>(0.01)                         | 16.72<br>(0.002) |                 |                 |
| 1998-00 | 39.67<br>(0.04)                       | 39.58<br>(0.01) | 18.41<br>(0.01)                         | 17.00<br>(0.002) |                 |                 |
| 2003-05 | 39.10<br>(0.04)                       | 40.38<br>(0.02) | 18.88<br>(0.01)                         | 17.29<br>(0.003) |                 |                 |
| 2008-10 | 38.44<br>(0.04)                       | 41.15<br>(0.02) | 19.28<br>(0.01)                         | 17.51<br>(0.003) |                 |                 |
| 2013-14 | 39.21<br>(0.04)                       | 41.30<br>(0.02) | 19.60<br>(0.02)                         | 17.82<br>(0.004) |                 |                 |
|         | <i>Immigrants by region of origin</i> |                 | <i>Immigrants by region of origin</i>   |                  |                 |                 |
|         | EU14                                  | A8              | non-EU                                  | EU14             | A8              | non-EU          |
| 1993-95 | 41.59<br>(0.09)                       | 46.29<br>(0.41) | 38.84<br>(0.05)                         | 17.29<br>(0.02)  | 19.58<br>(0.11) | 18.13<br>(0.02) |
| 1998-00 | 40.65<br>(0.09)                       | 38.89<br>(0.39) | 39.33<br>(0.05)                         | 18.15<br>(0.02)  | 19.77<br>(0.10) | 18.48<br>(0.02) |
| 2003-05 | 40.67<br>(0.09)                       | 32.73<br>(0.21) | 38.92<br>(0.05)                         | 18.78<br>(0.03)  | 19.92<br>(0.06) | 18.87<br>(0.02) |
| 2008-10 | 40.42<br>(0.10)                       | 31.43<br>(0.08) | 39.09<br>(0.04)                         | 19.07<br>(0.03)  | 19.97<br>(0.03) | 19.21<br>(0.02) |
| 2013-14 | 40.65<br>(0.12)                       | 33.70<br>(0.09) | 40.06<br>(0.05)                         | 19.81<br>(0.04)  | 19.93<br>(0.03) | 19.49<br>(0.02) |

Standard errors in parentheses.

Source: Quarterly LFS

**Figure 1.2: Trend in age distributions of immigrants and natives**

## 1.2 The labour market outcomes of immigrants

Despite their educational attainment advantage and younger age, immigrants generally experience lower employment probabilities in comparison to natives, with the exception of A8 immigrants in recent years. Looking at Figure 1.3, we observe that the employment probability amongst immigrants has consistently

been below the employment probability amongst natives over the last two decades. The gap in employment probabilities between the two groups has, nevertheless, been shrinking over the years: while the gap in employment probabilities between immigrants and natives was close to 10 percentage points around the mid 1990's, it dropped to about 5 percentage points in 2014. A closer look at the figure reveals two key observations: first the trend in the employment probability of immigrants with longer stay in the UK is largely similar to the trend in the employment probability of natives, and second there was a sharp increase in the employment probability of A8 immigrants from the mid-2000's, overtaking that of all other subgroups reported in the figure.

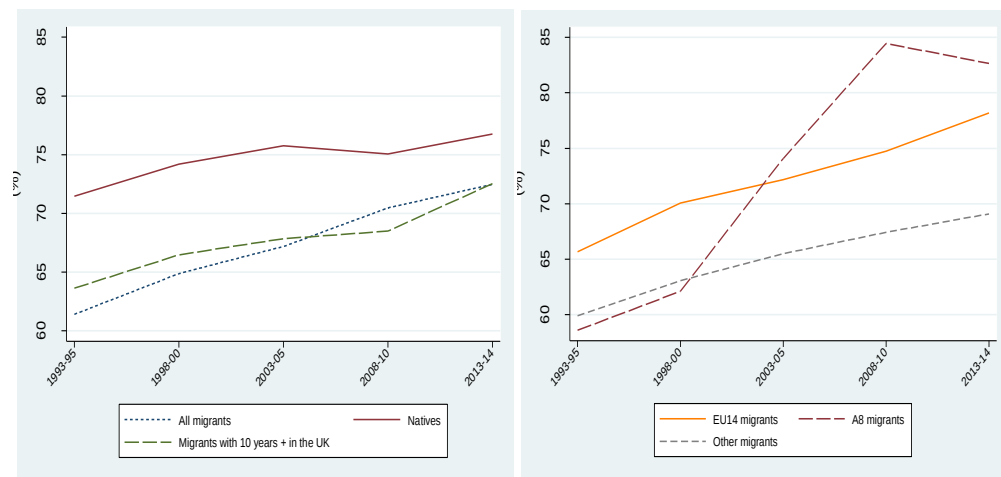
As with employment probabilities, the occupational distribution of immigrants with lengthier stay in the UK follows a similar pattern as that of natives. These two groups have much higher shares of employment in the upper level of the occupational ladder; close to half are in a managerial, professional or associate professional occupation. As reported in Table 1.2, on the other hand, only about one third of employed immigrants with less than 10 years of stay in the UK are in these occupations. Elementary occupations comprise the largest share of employment amongst them, with 23.3 per cent of those in employment. In comparison, only 10.5 per cent of employed immigrants with 10 or more years of stay in the UK are in such occupations. Amongst this group of immigrants and natives, professional occupations comprise the largest share of employment. Such observations, in combination with what we saw in the case of employment probabilities, tentatively suggest that an immigrant may have better labour market outcomes in the host country as their stay in the country lengthens. As suggested by Chiswick (1978), immigrants acquire human capital specific to the host country with time spent in the country, which could lead to better labour market outcomes.

Table 1.3 presents occupational distributions across groups of immigrants as of late. There is a clear contrasts in these distributions: while over 30 per cent of A8 immigrants are employed in an elementary occupation, a little less than 10 per cent of EU14 immigrants are employed in these occupations. The corresponding share is about 13 per cent amongst non-EU immigrants. Managerial and professional occupations accounts for 42 per cent of employment amongst EU14

immigrants, but just over 10 per cent amongst A8 immigrants. For non-EU immigrants, the corresponding share is 36 per cent.

Table 1.1 to Table 1.3 and Figure 1.3 point out three key observations: 1) individual characteristics such as age and educational attainment levels differ between immigrants and natives; 2) subgroups of immigrants differ in their characteristics as well as in their labour market outcomes over time and across subgroups; and 3) A8 immigrants exhibit a distinct pattern of employment rate and occupational distribution.

**Figure 1.3: Employment probability over time**



Source: Quarterly LFS, individuals of age 25-64 only

**Table 1.2: Occupational distributions of immigrants and natives (in %), LFS (2013-2014)**

|                                    | Immigrants |                                   |                              | Natives |
|------------------------------------|------------|-----------------------------------|------------------------------|---------|
|                                    | All        | With less than 10 years in the UK | With over 10 years in the UK |         |
| Managerial                         | 9.63       | 6.13                              | 11.80                        | 11.42   |
| Professional                       | 23.35      | 18.85                             | 26.08                        | 21.34   |
| Associate Professional & Technical | 10.73      | 9.02                              | 11.78                        | 14.89   |
| Administrative & Secretarial       | 7.80       | 6.37                              | 8.72                         | 11.70   |
| Skilled Trades                     | 9.02       | 9.76                              | 8.55                         | 10.89   |
| Personal services                  | 9.94       | 9.99                              | 9.92                         | 9.10    |
| Sales & Customer Service           | 5.62       | 5.74                              | 5.52                         | 6.31    |
| Process, Plant & Machine Ops.      | 8.57       | 10.80                             | 7.17                         | 6.33    |
| Elementary                         | 15.34      | 23.33                             | 10.46                        | 8.00    |

Source: Quarterly LFS; individuals of age 25-64 only

**Table 1.3: Occupational distributions by group of immigrants (in %), LFS (2013-2014)**

|                                    | <i>EU14 immigrants</i> | <i>A8 immigrants</i> | <i>non-EU immigrants</i> |
|------------------------------------|------------------------|----------------------|--------------------------|
| Managerial                         | 11.54                  | 3.9                  | 10.55                    |
| Professional                       | 30.47                  | 6.4                  | 25.72                    |
| Associate Professional & Technical | 16.32                  | 4.9                  | 10.74                    |
| Administrative & Secretarial       | 9.25                   | 6.11                 | 7.85                     |
| Skilled Trades                     | 6.5                    | 15.45                | 8.08                     |
| Personal services                  | 8.11                   | 8.45                 | 10.78                    |
| Sales & Customer Service           | 4.53                   | 4.11                 | 6.26                     |
| Process, Plant & Machine Ops.      | 3.99                   | 18.25                | 7.35                     |
| Elementary                         | 9.29                   | 32.44                | 12.67                    |

*Source: Quarterly LFS; individuals of age 25-64 only*

Understanding how immigrants fare in the host country's labour market has been of particular interest to both academics and policy makers:

“The issue of immigration and the integration of immigrants and their children are high on the policy agenda of the EU and OECD countries, both from an economic and a social standpoint. The active participation of immigrants and their children in the labour market and, more generally, in the public life is vital for ensuring social cohesion in the host country and the ability of migrants to function as autonomous, productive and successful, self-realised citizens. This is critical for facilitating their acceptance by the host-country population.” [OECD 2015. “Indicators of Immigrants Integration, Settling In” p.9]

The need for better understanding of the assimilation of immigrants and their children in the host country was further brought to the policy sphere in recent rounds of general elections across OECD countries. And one cannot, but notice, the re-emergence of its importance in recent debates about the Syrian migrant crisis across Europe.

Chiswick's 1978 seminal paper on the earnings growth of male immigrants in the USA led to a growing literature on the labour market outcomes of immigrants in the host country. The subject was in the academic sphere 170 years ago in Britain, however; Engels (1845) provided one of the first accounts of the labour market outcomes and living conditions of (Irish) immigrants in Britain in his work entitled “The Condition of the Working Class in England”. In recent years, a large literature has examined various labour market outcomes of UK immigrants, relative to their native counterparts. While a number of these studies have focused on relative earnings (Chiswick, 1980; Bell, 1997; Drinkwater et al,

2009; and Gagliardi and Lemos, 2015), others have examined a number of other labour market outcomes such as labour market participation, employment probabilities, self-employment, industries and occupations of employment (Dustmann and Faber, 2005; Clark and Drinkwater, 2008; Algan et al, 2010; and Longhi and Rokicka, 2012). A general finding from these studies is that immigrants have, on average, worse labour market outcomes in comparison to their native counterparts. When immigrants are grouped according to their countries of origin, however, the findings reveal that immigrants from countries with comparable income levels as the UK (or similar economic structure) tend to have better relative labour market outcomes than other groups of immigrants.

Part of the inferior labour market outcomes of immigrants, relative to natives, has been attributed to their lack of human capital specific to the host country such as language skills, education/training, experience (see Chiswick (1978) and Dustmann and Fabbri (2003) for example). While one would not expect the children born to immigrants in the host country (i.e. second generation immigrants) to lack these skills, empirical evidence points out some correlation between their labour market outcomes and those of their parents (see Borjas (1994)). A tranche of the literature on the assimilation of immigrants in the UK goes further to examine the relative labour market outcomes of second generation immigrants (Heath and Li, 2008; Algan et al, 2010; and Dustmann and Theodolopoulos, 2010). A key objective of this literature is to examine the presence (or absence) of intergenerational mobility in the labour market outcomes of immigrants by comparing these outcomes amongst first generation and second generation immigrants to the same outcomes amongst natives. One common generalisation from these studies, however, has been to categorise all UK-born individuals of non-White ethnicity as second generation immigrants, principally because of data limitations. In light of the long history of immigration of some ethnic groups in the UK, such an approach would likely lead to classification errors, which may have severe implications for inference of intergenerational mobility in the labour market outcomes of interest.

Some authors have suggested that the differences/comparability in the labour market outcomes of groups of immigrants, relative to natives, can be attributed to differences in the “quality” of immigrants from the respective groups (examples include Borjas (1987) and Mattoo et al (2008)). Borjas (1987), for

instance, argues that immigrants would differ in skills depending on the comparability of income levels and differences in inequality levels between the host country and the home country of the immigrant. Differences in the labour market outcomes of immigrants may also be desirable for the host country, particularly in terms of occupations/sectors of employment, as pointed out by Green (1999).

While the host country might be able to influence these outcomes for some groups of immigrants, it may have limited influence on them for other immigrants. For example, before the 1962 Commonwealth Immigration Act, Commonwealth citizens had free access to the UK's labour market; the same citizens later required a work permit to enter the UK's labour market when the Act was implemented. Today, nationals from countries that are member states of the European Union can freely enter the UK and take up employment in any occupation; immigrants from other parts of the world, on the other hand, generally require a work permit to enter the UK for employment purposes. Because these permits are often tailored toward particular occupations, the occupational outcomes of immigrants may, therefore, result from a combination of the host country's immigration policy and the characteristics of the source country of the immigrant. The literature provides very scarce evidence on the role of immigration policy on the occupational outcomes of immigrants, an exception being Green and Green (1995) in the context of Canada.

Beside immigration policy, social capital – i.e. family and friends – may play a role in the occupational outcomes of immigrants (Patel and Vella, 2013). The importance of family and friends in labour market outcomes has been highlighted by a large and burgeoning literature (examples include Rees (1966), Granovetter (1973), Ioannides and Loury (2004), Wahba and Zenou (2005), Bayer et al (2008) and Cappellari and Tatsiramos (2015)). Empirical evidence on the occupational (and economy wide) implications of networks of social contacts are examined in Bentolila et al (2010) in the context of Europe and the USA. Patel and Vella (2013) provide a more direct link between the occupational composition of the network and the occupational outcome of an immigrant job seeker in the context of the USA. Despite the benefits associated with the use of networks of contacts in the labour market, these studies suggest that they could have non-positive implications. While Bentolila et al (2010) show that they can lead to lower

productivity levels in the economy, Patel and Vella (2013) argue they may explain the concentration of certain groups of immigrants in particular occupations in metropolitan areas in the USA.

In the case of immigrants, these networks have commonly been defined in empirical studies as co-ethnics of the immigrant in the host country (Munshi, 2003; Beaman, 2012; and Patel and Vella, 2010). While the immediate source of information to recent immigrants is likely to be co-ethnics in the host country (Munshi, 2003; and Kanas et al, 2012), an immigrant could potentially broaden their circle of friends/acquaintances with time spent in the host country. The empirical evidence on how the occupational composition of those in employment in the network affects the occupational outcomes of the job seekers is rather limited. Moreover, in the context of immigrants, we know little about how these effects vary with their length of stay in the host country.

***Brief summary of gaps in the literature and our contribution.***

Immigration policy is used in several OECD countries as a mechanism to influence labour supply in certain occupations. In the UK, for example, a special committee regularly provides the government with a list of occupations in which there is a shortage of workers, and there then is an effort to fill these shortages with immigrant workers. While selection systems such as this are becoming increasingly used by many countries, empirical evidence on the role of immigration policy on the occupational outcomes of immigrants is scarce. Yet the implementation of these policies are costly to all parties involved. At the same time, although there has been a large number of studies looking at the determinants of occupational success of immigrants once they are in the host country, only very few have looked at the influence of social capital in these outcomes despite the widely held knowledge that immigrants, particularly those who are new in the host country, rely on fellow co-ethnics as a source of job information. As such, Chapter 2 and Chapter 3 focus on the occupational outcomes of immigrants, examining the roles of immigration policy and that of social capital. Chapter 2 will examine how important the role of immigration policy is for the occupational outcomes of immigrants. Chapter 3 focuses on the role social capital in these occupational outcomes.

Chapter 2 contributes to the literature by teasing out the role of immigration policy on the occupational outcomes of immigrants from other factors that may affect these outcomes. Chapter 3 contributes to the scarce literature on the link between the occupational composition of a job seeker's network of social contacts and their own occupational outcome, using direct information on whether a job seeker found employment through referral or other search methods. Of particular interest here is whether these effects vary with the length of stay of the immigrant in the UK.

Chapter 4 takes an intergenerational perspective on the assimilation of UK immigrants and departs from looking at occupational outcomes. It compares the labour market outcomes of different groups of first generation immigrants and second generation immigrants, respectively, to the outcomes of natives. A key contribution of this essay is to show the extent of the potential statistical bias resulting from using the common proxy measure of nativity status when examining the intergenerational assimilation of immigrants in the UK's labour market.

### **1.3 Outline of the thesis**

The thesis starts with Chapter 2, which evaluates the role of immigration policy on the occupational outcomes of immigrants. On the 1<sup>st</sup> of May 2004, nationals of the A8 countries became (quasi) free from immigration restrictions to live and work in the UK. Immigrants from EU14 countries have been free from these restrictions since at least 1995. Immigrants from non-EU countries continue to face immigration restrictions in the UK. The chapter uses the exogenous annulment of immigration restrictions faced by A8 immigrants in the UK in a difference in differences framework to isolate the role of immigration policy on the occupational outcomes of immigrants. The findings suggest that although there was a sharp increase in the share of A8 immigrants who entered the UK after 2004 in elementary occupations, this was not entirely the result of the change in immigration policy. It was, at least in part, a result of the fact that A8 immigrants who entered the UK after 2004 had shorter duration of stay in the country. Looking at a sample of immigrants who entered the UK shortly before and after 2004, the findings suggest that the change in the UK's 2004

immigration policy led only to a slight increase, although not statistically robust, in the odds of observing A8 immigrants who entered the UK after 2004 in elementary occupations relative to professional occupations. The findings from this chapter could be of interest to policy makers, in particular those who are interested in freedom of movement of labour agreements. Additionally, A8 immigrants have a different occupational distribution to natives, which may imply that they are not in direct competition with natives in the labour market.

Comparing immigrants with different lengths of stay in the UK, we find that A8 immigrants with longer stay in the country have better occupational outcomes. This is in line with findings that immigrants move up the occupational ladder with time spent in the host country. Accounting for income levels at the source country of the immigrant in the estimation sharply reduces differences in the occupational outcomes of A8 immigrants and those of UK-born individuals. For instance, the odds of being in higher level occupations increase sharply amongst immigrants with over 10 years of stay in the UK when compared with UK-born individuals. These findings could be interest to policy makers, in particular those in OECD countries where the assimilation of immigrants is high on the political agenda. While this chapter does not examine factors that lead to better labour market outcomes for immigrants with longer stay in the host country, known examples of these include acquisition of language skills and targeted training for other skills required in the host labour market. Therefore, strengthening or encouraging policies that can facilitate or accelerate the acquisition of these factors could lead to better labour market outcomes amongst immigrants.

Chapter 3 examines to what extent the network of social contacts of an immigrant job seeker affects their occupational outcome. It is now a stylised fact that networks of social contacts are widely used in the labour market, a principle reason being that they reduce asymmetry of information between employers and prospective employees, and thereby lower search costs (Rees, 1966). An immigrant's network of social contacts is defined as their co-ethnics in the host country, as it is commonly the case in the literature. This chapter is particularly interested in whether the influence of such networks on the occupational outcomes of job seekers varies across immigrants with different lengths of stay in the host country as this could entail the presence (or absence) of some form of "social assimilation".

We are able to directly observe in the data whether employment was obtained through referral or not. Therefore, unlike previous studies, we can first empirically examine the influence of a network of fellow co-ethnics on the occupational outcomes of immigrants with different durations of stay in the host country. We then examine, as in Patel and Vella (2013), whether a job seeker is likely to choose the most popular occupation amongst members of the network who are in a position to provide referrals or credible job information. The empirical analysis concludes by looking at the relationship between the size of one's network in a given occupation and the likelihood of ending up in that particular occupation. The evidence suggests that the occupational composition of fellow co-ethnics, the network, has less pronounced effects on migrants with lengthier durations of stay in the UK.

The findings from this chapter would appear to suggest that co-ethnics may form the initial pool of social contacts of a new immigrant, but they become “socially assimilated” by broadening their network of social contacts with time spent in the host country. It seems from the data that these broader networks do not lead to receiving less job information for more established migrants. As such, new immigrants would likely be more exposed to negative employment shocks amongst compatriots, compared to more established immigrants. Encouraging familiarity with alternative job information agencies/resources such as job centres amongst new immigrants may alleviate these concerns.

Chapter 4 looks at the implications of classification errors for parameter estimates of interest when the regressor is a binary variable. We illustrate these implications by examining the labour market outcomes of immigrants in an intergenerational context. Studies on the assimilation of immigrants in the UK often make no clear distinction between the nativity status of a UK-born individual and his/her ethnicity (see for example Algan et al (2010) and Heath et al (2008)). This is because of lack of information on parental country of birth in the LFS, the source of the data used therein. When using the LFS, researchers have often classified individuals of non-White ethnicities born in the UK as second generation immigrants although their parents could potentially be children or grandchildren of UK-born individuals. The UK General Household Survey (GHS), which contains information on parental country of birth of respondents, suggests that over 10 per cent of the UK-born ethnic minority

population have parents who themselves were born in the UK. Hence, the categorisation adopted when using LFS data would very likely lead to individuals being assigned the incorrect nativity status. These classification errors, even at a small magnitude, can have serious consequences for the identification of the regression coefficients (Kreider, 2010).

Using data from the GHS, one can obtain nativity status by 1) using information on parental country of birth and 2) following the categorisation adopted when using LFS data. Our findings suggest that errors caused by wrongly classifying individuals as second generation immigrant lead to estimation bias. This bias may over or underestimate coefficients by over 30 per cent. It is also possible to examine the extent of these biases using LFS data because one can introduce a small percentage of classification errors in the sample. As such, we randomly allow 1 per cent of non-White individuals born in the UK to be misclassified in the LFS data. Comparison of the resulting estimates to those obtained from assuming zero error of classification corroborates our earlier findings.

## **Chapter 2: The role of immigration policy on the occupational outcomes of immigrants**

## 2.1 Introduction

On the 1st of May 2004 the United Kingdom (UK) allowed nationals of the A8 countries (i.e. The Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia) virtually free access to its labour market. From this date onward, immigrants from these countries no longer required a visa to either enter the UK or take up employment in the country. This led to an increase in the share of working age individuals from A8 countries in the UK labour force and a shift in their occupational distribution in comparison to other immigrants as well as natives. Around 40 per cent of A8 immigrants who entered the UK from 2004 onward were employed in an “elementary occupation”. The corresponding shares were less than 20 per cent for immigrants from non-A8 countries. In comparison, no more than 21 per cent of immigrants from A8 countries who entered the UK before 2004 were employed in these occupations. Despite this large change in occupations, the educational attainment and age at arrival in the UK of A8 immigrants did not change much before and after 2004, implying that the (observable) skill composition of these individuals did not change greatly. This suggests the shift in the occupational distribution of A8 immigration may have been a result of the change in the pre-existing immigration restrictions faced by A8 immigrants in the UK.

The existing literature posits that the inferior labour market outcomes of groups of immigrants arise from lower skill levels (see for example Borjas (1985, 1987) and Mattoo et al (2008)). Looking at immigrants in the USA, Mattoo et al (2008) argue that one may observe immigrants with similar observables at different points of the occupation ladder because quality of education may vary across the origin countries of immigrants, implying unobserved differences in human capital levels. It has also been suggested that the host country’s policies may play a role (Green and Green, 1995). Green (1999) further suggests that differences in the occupational distribution of immigrants and natives might be desirable to the host country as this could indicate that immigrants are filling supply gaps in the labour market.

Several OECD countries adopt immigration policies whereby some immigrants are granted permission to enter the country in order to fill shortages in certain occupations or industries. In the UK, for example, while the requirement of a work permit would dictate the occupational outcomes of immigrants from some countries, immigrants from member states of the European Union are free to enter the country and take up employment in any occupation. It is therefore possible that the observed differences in the occupational distributions of groups of immigrants may result from a combination of the underlying heterogeneity in the source countries of immigrants and the host country's immigration policy. The question we ask in this chapter is how much of the occupational outcome of an immigrant can be attributed to the host country's immigration policy and how much can be attributed to characteristics linked to the source country?

Using a difference in differences design, we exploit the exogenous annulment of immigration restrictions faced by A8 immigrants in the UK before 2004 to separate the role of immigration policy from the role of the source country. Immigrants from all pre-existing member states of the EU have been free from immigration restrictions in the UK since at least 1995, and non-EU immigrants remain under immigration restrictions similar to those faced by A8 immigrants before May 2004 – i.e. to enter, to work or remain in the UK. Immigrants from pre-existing member states of the EU and non-EU immigrants therefore provide natural counterfactual candidates with which to estimate the effect of the policy change on occupational outcomes. Looking at the occupational outcomes of a sample of recent immigrants who entered the UK shortly before and after the change in policy ensures that we avoid potential structural changes in their source countries that could lead to a sudden change in the (unobserved) skill composition of their nationals.

The closest study to ours, to the best of our knowledge, is Green and Green (1995) who examine the role of immigration policy on the occupational outcomes of immigrants using the introduction of the point based system in Canada in 1967. Their findings suggest the change in policy affected the occupational composition of immigrants. First, however, it is possible that this resulted from a change in the composition of source countries of immigrants rather than a change in the occupational distribution of immigrants from a given country (Borjas, 1993). Second, they do not observe actual occupations and

instead use future intended occupations at point of entry. Yet, as shown by Chiswick et al (2005), immigrants initially downgrade their occupations once in the host country's labour market and move up the occupation ladder with time spent in the country. Green (1999) also find evidence of upward occupational mobility amongst immigrants as their stay in Canada lengthens.

The contribution of the present chapter is that we are able to tease out the role of immigration policy amongst other factors that affect the occupational outcomes of immigrants using a quasi-natural experiment. Our findings suggest that the change in the UK's 2004 immigration policy led only to a slight increase, although not statistically robust, in the odds of observing A8 immigrants who entered the UK after 2004 in elementary occupations relative to professional occupations. Measures of (observed) human capital also did not change much amongst cohorts of A8 immigrants entering the country shortly before and after 2004, suggesting not much changed in the skill composition of A8 immigrants following the annulment of immigration restrictions. The large shift in occupations after 2004 appears to, at least in part, result from immigrants who entered the UK after 2004 having shorter duration of stay in UK. Consistent with findings from studies in other settings, we find that A8 immigrants with longer durations of stay in the UK have better occupational outcomes. Accounting for income level of the source country of the immigrant in the estimation sharply reduces differences in the occupational outcomes of A8 immigrants and those of UK-born individuals.

In what follows, we start with a brief review of the UK's immigration policy in section 2 and present a review of the literature in section 3. Section 4 outlines the hypotheses motivating our empirical analysis. Section 5 is devoted to the data and Section 6 outlines our methodology. We present our estimation strategy and results in section 7. Section 8 provides some discussions and section 9 concludes the chapter.

## **2.2 Institutional context**

Up until the 1970s, the UK mostly relied on the Commonwealth for its supply of labour to compensate for domestic shortages after the war. Commonwealth

citizen had free access to the UK's labour market until 1962; they filled the labour shortage the country experienced post-1945, especially in sectors such as health and transport. The 1962 Commonwealth Immigration Act saw the beginning of a more restrictive immigration policy with the introduction of employment "vouchers", limiting entry for those without a job offer or without "skills in short supply" in the country. As noted by Spencer (2011), "while in 1945 Commonwealth citizens had free access to live and work in the UK, by the 1970s their position was largely reduced to that of aliens. The fortune of Europe's citizens was the mirror opposite: aliens in 1945 but enjoying free movement within the European Economic Community (EEC) three decades later".

A major change in the UK's immigration policy was the Immigration Act 1971, which came into force in 1973, the year the UK became a member state of the EEC. The Act introduced the concept of the 'right of abode', defining persons exempt from immigration control in the UK. An individual with the right of abode does not need permission to enter, to work or to remain in the UK. The British Nationality Act 1981 redefined those with the right of abode in the UK, and as of January 1983 those with this right were: a citizen of the UK and Colonies (CUKC) born, adopted, naturalised or registered in the UK or the Islands (the Channel Islands and the Isle of Man); a CUKC born to or adopted by a parent, who at the time of birth, had the right of abode; a CUKC who was ordinarily resident in the UK for any continuous period of 5 years before 31 December 1982; a Commonwealth citizens with a parent or adoptive parent who, at the time of birth/adoption, was a CUKC by birth in the UK; and a female Commonwealth citizen/CUKC woman who is, or has been married to a man with right of abode before 31 December 1982<sup>2</sup>. The right of abode is only acquired by becoming a British citizen from January 1983. Nationals of EU member states also have the right of abode in the UK<sup>3</sup>, obtained from the freedom of movement of goods, workers and family members between member

---

<sup>2</sup> Further details on this can be found on: <https://www.gov.uk/government/publications/right-of-abode-roa/right-of-abode-roa#roa1-what-is-the-right-of-abode>

<sup>3</sup> Those from Romania and Bulgaria, the more recent members, still faced some restriction on work in the UK at the time of writing

states of the EU, founded in 1993. Non-EU nationals therefore faced immigration restrictions in the UK<sup>4</sup>.

Efforts by successive governments to align immigration with the need of the economy saw the conception of several restrictive measures, tailoring entry into the UK for employment purposes towards particular occupations. As per the Statement of Change in Immigration Rule (HC 395), by October 1994 non-EU nationals entering or already in the UK with prime purpose of undertaking some form of employment could be separated into two groups<sup>5</sup>:

- 1) Those seeking to enter or remain in the UK in an 'Au Pair' Placement, as a Working Holidaymaker, or for Training or Work Experience – this included seasonal workers at agricultural camps, teachers and language assistants coming to the UK under approved exchange schemes. Full-time employment was not a requirement for the first category. In fact, Au Pairs could only work for a maximum of 5 hours per day with 2 free days per week. Most leaves issued in this category were for a maximum period of 1 year, with the exception that those coming for training purposes could obtain a leave of 3 years and Au Pairs could obtain a leave of 2 years. Anyone could later apply for an extension to the duration of their leave; and
- 2) Those seeking to enter or remain in the UK for employment. Individuals primarily seeking to enter or remain in the UK for employment purpose were to hold a work permit or a form of leave in one of the following: Representative of overseas newspapers, news agencies and broadcasting organisations; Representative of overseas firms which have no branch, subsidiary or other representative in the UK (sole representatives); Private servants in diplomatic households; Overseas government employees; Minister of religion, missionaries and members of religious orders; Airport-based operational ground staff of overseas-owned airlines; Business persons, Inventors, Artists; and those given exception on the ground of UK ancestry – the individual had to be a commonwealth citizen, aged 17+ and able to prove

---

<sup>4</sup> They may cease to be under immigration control upon naturalisation as a British citizen or when granted indefinite leave to remain in the UK, either of which requires that the individual has spent a good number of years in the UK.

<sup>5</sup> By law, the UK has to open its border to asylum seekers but these individuals are often not allowed to work while awaiting decision on their application, at least in the first 6-12 months. Family members of those granted permission to work in the UK may also be entitled to work in the UK.

that one of his grandparents was born in the UK and Islands, amongst other things. A work permit and exceptional leave on the ground of UK ancestry could have a maximum duration of 4 years, but most other leaves had a maximum duration of 1 year.

As shown in Table 2.1, work permits were the predominant form of leave granted to enter or remain in the UK for employment purposes up until 2008, the year of the introduction of the Point Based System (PBS). Employers were required to provide extensive evidence for the need to recruit outside the EU. They had to supply copies of all advertisements of the vacancy they wished to recruit a non-EU national for, showing the name and date of publication. The advertisement needed to identify a specific post and “must have been placed within the last 6 months before the application being made in a quality newspaper with national or EU circulation, or in a trade journal.” [Supperstone et al (1994), and Macdonald and Webber (2005)].

**Table 2.1: Passenger admitted by purpose of journey, 1998 to 2003 (excludes EU nationals)**

|                                    | 1998          | 1999          | 2000          | 2001          | 2002          | 2003          |
|------------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <b>Work permit holders</b>         | <b>48,180</b> | <b>53,410</b> | <b>66,945</b> | <b>81,065</b> | <b>85,620</b> | <b>81,400</b> |
| Working holidaymakers              | 40,800        | 45,800        | 38,400        | 35,800        | 41,700        | -             |
| Au pairs                           | 15,900        | 14,900        | 12,900        | 12,000        | 12,800        | 15,300        |
| UK ancestry                        | 10,200        | 11,900        | 10,900        | 10,700        | 10,400        | 9,200         |
| <i>Dependent of permit holders</i> | <i>20,205</i> | <i>22,600</i> | <i>24,880</i> | <i>27,760</i> | <i>34,495</i> | <i>37,800</i> |

Source: 1998 to 2002 figures are from *Control of Immigration: Statistics United Kingdom, 2002*

<http://webarchive.nationalarchives.gov.uk/20110218135832/rds.homeoffice.gov.uk/rds/pdfs2/hosb1003.pdf>

2003 figures are from *Control of Immigration: Statistics United Kingdom, 2003*

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/251055/6363.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/251055/6363.pdf)

**Table 2.1 continues... Passenger admitted by purpose of journey, 2004 to 2008 (excludes EU nationals)**

|                                    | 2004          | 2005          | 2006          | 2007          | 2008          |
|------------------------------------|---------------|---------------|---------------|---------------|---------------|
| <b>Work permit holders</b>         | <b>82,700</b> | <b>91,500</b> | <b>96,600</b> | <b>89,200</b> | <b>81,900</b> |
| Other work categories              | 89,900        | 81,000        | 68,800        | 57,100        | 46,900        |
| Domestic Workers                   | 10,400        | 10,100        | 12,500        | 11,300        | 11,500        |
| UK ancestry                        | 7,700         | 8,260         | 8,490         | 7,450         | 6,700         |
| <i>Dependent of permit holders</i> | <i>42,000</i> | <i>45,900</i> | <i>49,000</i> | <i>38,900</i> | <i>32,600</i> |

Source: Table ad.03.w: *Passengers given leave to enter the United Kingdom in employment categories by country of nationality; figures are rounded to the nearest thousand. 1,060 permits were granted through the PBS in 2008*

Note: the difference in classes of entry result from the categorisation available from the source of data

Additionally, for employers wishing to employ a non-EU national who is a keyworker, e.g. those with specific language skills or cultural knowledge, the advert had to be available in all EU countries. Evidence required for the application of visa includes:

1. Special skills, experience or qualifications which the worker has making him uniquely qualified to do the job;
2. The way in which the job of workers and the success of the business depend on the recruitment of the person; and
3. The relevant knowledge of language or cultural skills which are not readily available in the EU, together with reasons why these skills are essential for the job.

Employers were not required to provide as much evidence for occupations on the shortage list. The shortage list comprises occupations where the authorities estimate there is not enough domestic workers (or workers from the EU) to meet demand<sup>6</sup> – the responsibility to assess which occupations would be on this list was assigned to an independent body in 2007, the Migration Advisory Committee (MAC)<sup>7</sup>.

In May 2004, nationals of the A8 countries became free from any restriction to enter or live in the UK when A8 countries became member states of the EU. They obtained the freedom to take up employment in any occupation, without any form of visa requirements from the UK government, except that some were required to register under the Worker Registration Scheme (WRS) with the UK Border Agency (UKBA); those who were already resident of the UK did not require registering under the WRS. As noted by MAC, "... The WRS was introduced in the UK as a transitional measure that monitors A8 nationals' access to the UK labour market. So long as they register their employment with the UKBA under the scheme, A8 nationals are free, as are nationals of other EU countries, to work in any occupation or sector and for any employer. ...On receipt of the registration form, UKBA will carry out checks to ensure that an

---

<sup>6</sup> Work permits issued through the shortage occupation route accounted for a relatively small proportion of work permits – see MAC's September 2008 report: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/257245/shortageoccupationlistreport.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/257245/shortageoccupationlistreport.pdf)

<sup>7</sup> See <https://www.gov.uk/government/organisations/migration-advisory-committee/about> for the shortage lists of recent years.

employee is a genuine A8 national and that the employer is also genuine”. [Review of the UK’s transitional measures for nationals of member states that acceded to the European Union in 2004. Migration Advisory Committee Report, April 2009, paragraph 2.8 and 2.11]. More detailed requirements as prescribed by the law are presented in the appendix.

The cost of registration was minimal; it was £50 in 2005 (it went up to £70 in 2007, and was £90 in 2009), less than a third of weekly earnings from a full-time job which paid minimum wage at the time<sup>8</sup>. An employee from an A8 country had one month to register and was issued a registration card and a registration certificate which bore the name of their employer. The employee was required to obtain a new certificate (free of charge) when they change employer, but once they had continuously been in employment legally for 12 months, they were no longer required to apply for a new registration certificate whenever they changed employer.

We use this change in policy that occurred on 1<sup>st</sup> May 2004 to examine the role of the UK’s immigration policy on the occupational outcome of immigrants. As above-mentioned, immigrants from the pre-existing member states of the EU continue to be free from immigration restrictions in the UK, and non-EU immigrants still face immigration restrictions to enter, to work or remain in the UK. These two groups provide a natural counterfactual with which to estimate the effect of the policy change on occupational outcomes.

### 2.3 Related literature

The labour market performance of immigrants in the host country was brought to prominence by Chiswick (1978). The paper looks at the earning evolution of immigrants in the USA using 1970 census data. Chiswick finds that foreign-born adult males in the USA start with lower wages than their American-born counterparts but experience strong earnings growth: they catch up with natives within 10-15 years of arrival in the USA, and their earnings exceed those of the natives from then onward. In accordance with human capital theory, he posits that it takes time to acquire USA-specific labour market skills and the need for

---

<sup>8</sup> The minimum wage was £4.85 in October 2004 and £5.05 in 2005 for the 21+; full-time work is defined as 35+ hours/week.

immigrants to invest in their human capital depresses their wages in the short run. Once these skills are acquired, the wages of immigrants grow quite steeply. In a similar study on UK, Chiswick (1980) finds that duration of stay has little effects on the earnings of immigrants in Britain once one controls for experience. As argued by Borjas (1985), however, cross-section data studies on the assimilation of immigrants assume that immigrants from different cohorts are of the same quality and hence experience similar earnings growth with time spent in the host country. The author shows that the “time since migration effect”, often used to measure the assimilation of immigrants in the labour market, can be decomposed into a within-cohort effect (this gives a measure of the growth in earnings experienced by the same cohort across two time periods), and an across-cohort component (this measures the difference in earnings of two different cohorts with the same level of time spent in the destination country). Using data from the 1970 and 1980 censuses, Borjas shows that cohorts of immigrants differ in average “quality” and that the actual earnings growth associated with time spent in the host country is very marginal once this difference in quality is accounted for. This view is supported by Borjas (1987), who finds that disparities in return to education, income levels and income distributions between the host country and the sending countries determine the composition of immigrants and the subsequent differences in labour market outcomes of the respective groups of immigrants. Better labour market outcomes are associated with comparable income levels between the source and the host country.

Duration of stay in the host country and similarity of income levels between the source country and the destination country of immigrants may also affect occupational outcomes. Chiswick et al (2005) posit that the occupational attainment of immigrants follows a U shape: immigrants downgrade their occupations on arrival in the host country, but as their stay in country lengthens, they acquire skills specific to the host labour market and may attain their pre-emigration occupation levels. Using Australian longitudinal survey data with information on the immigrant’s occupation before and after entry into the country, they find evidence in support of their hypothesis. Factors such as similarity of labour market structure between the home and the receiving country, educational attainment and reasons for migrating also affect occupational attainment post emigration. Using the same survey and looking at

two separate cohorts, Mahuteau and Junankar (2008) find similar results. Green (1999) compare the occupational attainments of immigrants and natives in Canada using cross section data. The findings suggest that the occupational attainment of immigrants improves with length of stay in the country.

Using the 2001 Spanish population census, Amuedo-Dorantes and De la Rica (2007) examine the occupational attainment of immigrants relative to natives across 22 occupations for males and 17 occupations for females. Occupations are ranked according to their remuneration, and occupational outcomes are examined using an ordered logit estimator. They find no statistical difference between the occupational attainment of EU14 immigrants and that of Spanish nationals. Accounting for duration of stay in Spain, the occupational attainment differences for other non-EU14 European immigrants and those from Latin America declines, but the occupational attainment gap for African immigrants still remains. Another related study is Barrett and Duffy (2008) who evaluate the occupational outcome of different groups of immigrants in Ireland in comparison to natives. They find that immigrants are respectively 2 and 4 per cent more likely to be in unskilled and in middle-ranked occupations in comparison to natives, and they are 2 and 4 per cent less likely to be in professional and managerial occupations, respectively. They also group immigrants by time of arrival in Ireland and find while those who entered Ireland between 1995 and 1999 show higher probabilities of having better occupational attainments in comparison to natives, cohorts who entered after 1999 all have higher probabilities of being in lower level occupations. The occupational attainment of natives and nationals of countries with similar labour market structure as Ireland is not statistically different.

Barret and Duffy (2008) explicitly compare the occupational outcomes of immigrants from A8 countries to those of Irish nationals. This study is therefore of particular relevance to the present paper given that Ireland was one of the three EU member states that allowed A8 nationals to freely access its labour market from May 2004. The paper finds that A8 nationals have lower occupational attainments relative to Irish natives but there is no strong evidence of differences in occupational attainments between those who arrived in Ireland pre and post 2004, including those who arrived in the 1990's. They however look at a single year of cross section data, i.e. the 2005 Irish Quarterly National

Household Survey, shortly after May 2004. It is hence unclear whether the pre-existing Irish immigration restrictions and/or its labour market did not allow movement along the occupation ladder. One may also argue that earlier cohorts of immigrants may have had much worse occupational outcomes at entry than later cohorts, and these outcomes subsequently improved with time spent in Ireland.

In the context of the UK, Longhi and Rokicka (2012) look at characteristics and various labour market outcomes of EU immigrants in the country. Using LFS data over the period of 1997 to 2010, they find differences in the regional distribution, economic activity status, types of employment (i.e. temporary, self-employed, part-time or full-time employment) and industry of employment of A8 immigrants in the UK, before and after 2004. Because the outcomes of those who entered the UK before and after 2004 are compared (relative to the natives) regardless of their length of stay in the UK and their age at entry, these differences could arise from older immigrants having become more familiar with the host country's institutions with time spent in the UK, and therefore have closer outcomes to natives. For instance, comparing various cohorts of Polish immigrants in the UK, the findings of Drinkwater et al (2009) show that some older cohorts have higher earnings than younger and more recent cohorts. Similar findings are obtained for other A8 immigrants. However, in both cases, these conclusions become unclear when controls such as tenure with the firm, industry and occupations of employment are included in the regression: the estimated coefficients of interest fluctuate and most coefficients lose statistical significance.

A paper which looks at how changes in the host country immigration policy affect immigrants' occupational outcomes is Green and Green (1995). As far as we are aware, this study is the closest to ours. Green and Green find that the point based system introduced in Canada in the late 1960s led to a change in the (expected) occupational composition of the inflow of immigrants. Because occupations are not observed in the data, intended occupations provided at point of entrance are used as an alternative measure although these may not necessarily be the same. As pointed out earlier, immigrants tend to downgrade their occupations once in the host labour market, a phenomenon which could potentially be amplified by institutional constraints. Such case is shown in

Mahuteau and Junankar (2008), for example, where they examine the impact of lengthening residency duration for access to welfare eligibility on the labour market outcomes of immigrants in Australia. They find that while the change in policy shortened unemployment duration for those affected by it, it led to more initial downgrading of occupations amongst them.

Our study is, to the best of our knowledge, the first to look at how the complete relaxation of immigration policies affects immigrants' occupational outcomes. The novelty of the study is that we are able to separate the role of immigration policy on the occupational outcomes of immigrants from other contributing factors using an exogenous change in immigration policy. Countries to a large extent select immigrants based on observables that are often strong determinants of occupational outcomes themselves. Additionally, these immigration policies vary considerably according to the immigrant's place of birth and even sometimes across groups of individuals from the same country – i.e. economic migrants versus refugees for example. Moreover, the implementation of these immigration restrictions are costly the policy makers, employers and economic migrants. It is therefore of interest to policy makers what the occupational choices of immigrants would be in the absence of immigration constraints. The empirical evidence on the question is scarce. This essay contributes to filling the gap. The study may also be of interest to policy makers interested in freedom of labour agreements, and to EU policy makers with regards to countries applying for membership of the union.

#### **2.4 Ex-ante expectation of the implications of the relaxation of visa requirements**

In this section, we ask what economic theory would suggest about the implications of a complete relaxation of immigration policies that allows a group of immigrants free access to the UK's labour market.

In a series of influential articles (i.e. Borjas (1987, 1989, 1994)), Borjas provides theoretical insights on the determinants of migration flow between two countries and on the associated types of selections. Immigrants are either positively selected or negatively selected. Another type of selection presented is refugee sorting. The theoretical framework adopted in these articles is an extension (or a

more modern and clearer application) of Roy (1951). Roy's model suggests that in an economy with two sectors, workers self-select themselves into the sector that gives them the highest expected earnings. More recent studies that use Roy's model to examine patterns of international migration include Chiquiar and Hanson (2005) and Belot and Hatton (2012). Unlike Borjas, Chiquiar and Hanson present a scenario, followed by empirical evidence, where immigrants in a host countries are not drawn from one of the tails of the skill distribution in their home country. They coin this an "intermediary" scenario. Belot and Hatton (2012), on the other hand, present findings that shed light on why immigrants from low income countries that have higher return to education are positively selected.

The theoretical framework adopted in the above literature shows that immigration policy, through its impact on migration cost, can be used to influence both the size of the flow as well as the skill composition of incoming migrants. Empirically, while some authors have suggested that skill based selective immigration policies raise the skill composition of immigrants (Aydemir, 2011; and Czaika and Parsons, 2015) other empirical studies conclude these policies may not necessarily lead to much increase in the skill level of immigrants but may instead change the countries of origin composition of immigrants (Borjas, 1993; and Antecol et al, 2003). Because immigrant networks are an important determinant of the size of immigration flows as well as their skill compositions (Beine et al, 2011; and McKenzie and Rapoport, 2012), it has also been suggested that immigration policy can be used to alter the skill composition of these networks, which would subsequently affect the skill composition of immigrants in the long run (Bertoli and Rapoport, 2015).

In what follows, we derived hypotheses based on the theoretical framework presented in Borjas (1987). The selection type looked at here is selection on unobservables because occupations generally require a broader set of skills beyond schooling levels, and A8 immigrants entering the UK after 2004 are no longer subject to immigration controls. Earnings, and not occupations, are used to bring these insights to lights because it greatly simplifies the analysis. Nevertheless, their implications hold for the question we examine in this chapter,

not least because there exists a high level of correlation between earnings and occupations.

Suppose an individual in origin country  $o$  faces the following earning distribution:

$$\ln(w_o) = \mu_o + \varepsilon_o \quad (2.4.1)$$

If the individual emigrates to a destination country  $d$ , they face the following earning distribution:

$$\ln(w_d) = \mu_d + \varepsilon_d \quad (2.4.2)$$

$\varepsilon_o$  is normally distributed with mean 0 and variance  $\sigma_o^2$ , and similarly  $\varepsilon_d$  is normally distributed with mean 0 and variance  $\sigma_d^2$ .

As posits Borjas, the migration decision of an individual from country  $o$  to country  $d$  is determined by the sign of the index function (as per equation (3) in Borjas (1987)):

$$I = \ln[w_o/(w_d + c)] \approx (\mu_d - \mu_o - \pi) + (\varepsilon_d - \varepsilon_o) \quad (2.4.3)$$

$c$  is a mobility cost from country  $o$  to country  $d$ , and  $\pi$  measures a constant, time-equivalent cost of emigration, which can be expressed as  $c/w_o$ . It follows that the probability of an individual emigrating, i.e. the emigration rate, from country  $o$  to country  $d$  is as follows:

$$P = \Pr[v > -(\mu_d - \mu_o - \pi)] = 1 - \Phi(z) \quad (2.4.4)$$

where  $v = \varepsilon_d - \varepsilon_o$  and  $z = -(\mu_d - \mu_o - \pi)$ . From this equation it is straight forward to see that the probability of emigrating decreases in  $\mu_o$  and  $\pi$ , but increases in  $\mu_d$ . A direct implication is that country  $d$  can influence immigration flows from country  $o$  using immigration policy to alter the cost of migration: increasing (decreasing) migration cost would decrease (increase) migration flow from country  $o$  to country  $d$ .

In addition to providing insights into the size of migration flow between country  $o$  and country  $d$ , equation (2.4.4) informs the nature of selection in the decision to migrate. Under normality, it can be shown that, as suggests Borjas (1987):

$$E[\ln(w_o)|I > 0] = \mu_o + \frac{\sigma_o \sigma_d}{\sigma_v} \left( \rho - \frac{\sigma_o}{\sigma_d} \right) \left( \frac{\phi(z)}{1 - \Phi(z)} \right) \quad (2.4.5)$$

and

$$E[(\ln(w_d)|I > 0] = \mu_d + \frac{\sigma_o \sigma_d}{\sigma_v} \left( \frac{\sigma_d}{\sigma_o} - \rho \right) \left( \frac{\phi(z)}{1 - \Phi(z)} \right) \quad (2.4.6)$$

where  $\rho$  is the coefficient of correlation between  $\varepsilon_o$  and  $\varepsilon_d$  and  $\phi(z)$  is the density function of  $z$ .  $E[.]$  is an expectation term. The mechanism that guides the self-selection of immigrants can be seen from equations (2.4.5) and (2.4.6), in particular through the sign of  $\left( \rho - \frac{\sigma_o}{\sigma_d} \right)$  and  $\left( \frac{\sigma_d}{\sigma_o} - \rho \right)$ . Looking at equations (2.4.5) and (2.4.6), one sees that positive selection (negative selection) occurs when  $\frac{\sigma_o \sigma_d}{\sigma_v} \left( \frac{\sigma_d}{\sigma_o} - \rho \right)$  and  $\frac{\sigma_o \sigma_d}{\sigma_v} \left( \rho - \frac{\sigma_o}{\sigma_d} \right)$  are both positive (negative). Thus, we henceforth focus on one of them.

Assuming that  $Q_d = \frac{\sigma_o \sigma_d}{\sigma_v} \left( \frac{\sigma_d}{\sigma_o} - \rho \right) \left( \frac{\phi(z)}{1 - \Phi(z)} \right) = \gamma \lambda$ , as points out Borjas, the impact of changes in any variable  $x$  on the quality of immigrants from country  $o$  is given by:

$$\frac{\partial Q_d}{\partial x} = \lambda \frac{\partial \gamma}{\partial x} + \gamma \frac{\partial \lambda}{\partial x} \quad (2.4.7)$$

with  $\gamma = \frac{\sigma_o \sigma_d}{\sigma_v} \left( \frac{\sigma_d}{\sigma_o} - \rho \right)$  and  $\lambda = \left( \frac{\phi(z)}{1 - \Phi(z)} \right)$

while  $\lambda \frac{\partial \gamma}{\partial x}$  captures how a change in  $x$  induces incentives for different persons to emigrate from country  $o$  to country  $d$ ,  $\gamma \frac{\partial \lambda}{\partial x}$  captures how changes in  $x$  affect the size of migration flow from country  $o$  to country  $d$ . Borjas calls the former the “composition effect” as it relates to changes in the skill composition of incoming migrants, and he calls the latter the “scale” effect because it only affects the size of the flow (for any given selection type described earlier). Because migration cost, the instrument of immigration policy, does not enter  $\gamma$ , it follows from equation (2.4.7) that any change in migration cost would only affect the size of immigration flow and not the skill composition of immigrants.

One may argue that individuals in country  $o$  face heterogeneous migration costs such that  $\pi$  is a function of individual attributes (Borjas, 1999). In such case,

under the assumption that  $\pi$  is normally distributed with mean  $\mu_\pi$  and variance  $\sigma_\pi^2$ . The analogue of equation (2.4.4) will now be:

$$P = \Pr[v > -(\mu_d - \mu_o - \mu_\pi)] = 1 - \Phi(z) \quad (2.4.8)$$

with  $v = \varepsilon_d - \varepsilon_o - \varepsilon_\pi$ . It can be shown that equations that (2.4.5) and (2.4.6) now become, respectively (see Borjas, 1999):

$$E[\ln(w_o)|I > 0] = \mu_o + \frac{\sigma_o \sigma_d}{\sigma_v} \left[ \left( \rho - \frac{\sigma_o}{\sigma_d} \right) - \rho_{\pi o} \frac{\sigma_\pi}{\sigma_d} \right] \left( \frac{\phi(z)}{1 - \Phi(z)} \right) \quad (2.4.9)$$

and

$$E[(\ln(w_d)|I > 0] = \mu_d + \frac{\sigma_o \sigma_o}{\sigma_v} \left[ \left( \frac{\sigma_d}{\sigma_o} - \rho \right) - \rho_{\pi d} \frac{\sigma_\pi}{\sigma_o} \right] \left( \frac{\phi(z)}{1 - \Phi(z)} \right) \quad (2.4.10)$$

Here,  $\rho_{\pi d}$  and  $\rho_{\pi o}$  are correlation coefficients respectively between  $w_d$  and  $\pi$ , and between  $w_o$  and  $\pi$ . One sees that (2.4.5) and (2.4.6) are cases where  $\sigma_\pi = 0$  in (2.4.9) and (2.4.10). For  $\rho_{\pi d} = \rho_{\pi o} = 0$ , (2.4.9) and (2.4.10) also collapse into (2.4.5) and (2.4.6).

The components of (2.4.9) and (2.4.10) which now determine how changes in a given variable  $x$  induce incentives for different persons to migrate from country  $o$  to country  $d$ , i.e.  $\lambda$ , is now a function of the immigration policy of country  $d$ , unlike under a scenario of constant cost of migration. In other words, with heterogeneous migration cost, immigration policy can alter the skill composition as well as the size of migration flow. This will depend on the sign of  $\rho_{\pi d}$  and that of  $\rho_{\pi o}$ . We assume them to be negative, such that migration cost decreases with the skill level of the migrant from country  $o$  and also those with higher skill levels are likely to earn more in country  $d$  – we see this as a reasonable assumption. It therefore follows that  $\lambda$  would be increasing in  $\sigma_\pi$ , such that a reduction in  $\sigma_\pi$  decreases the skill composition of migrants from country  $o$ . This gives rise to our hypotheses:

### ***Hypotheses:***

- 1) Under constant migration cost, the annulment of visa policy does not affect the skill composition of new incoming A8 immigrants; and

- 2) The annulment of visa policy reduces variation in migration cost amongst A8 immigrants in the presence of heterogeneous migration cost. The reduction in  $\sigma_{\pi}$  lowers the skill composition of new incoming migrants. In other words, new incoming migrants would have lower occupational attainment levels. For small values of  $\rho_{\pi o}$ , however, the effect of the change in policy would be negligible.

In both cases, the annulment of visa policy leads to an increase in the size of migration flows given it reduces migration cost. The increase in the stock of A8 immigrants in the UK after 2004 is widely known and has been documented elsewhere (see Blanchflower and Shadforth (2009) for detailed exposition on the sharp increase in the stock of A8 immigrants in the UK after 2004). We will therefore focus on the implications for occupational outcomes in this chapter.

## 2.5 The data

The data we use comes from the UK Labour Force Survey (LFS). The LFS is the largest household survey of the UK population. Information is collected primarily on labour market issues, and the survey is used to calculate official measures of employment and unemployment in the UK. The LFS contains information on country of birth, labour market status, earnings, occupations, educational attainments, and other individual characteristics such as ethnicity, age, region of residence, etc.

The survey was first conducted in 1973 and every two years from then onward up until 1983. From 1984, it was conducted on an annual basis until 1991. Since 1992, it has been conducted on a quarterly basis. Each quarter corresponded to a season – i.e. winter, spring, summer or autumn – until May 2006 when it switched to calendar quarters: January-March, April-June, etc. The LFS covers those residing in private households, those residing in National Health Service (NHS) accommodation and young people living in student halls or residence during term-times. It is intended to be representative of the whole population of the UK.

The survey takes the form of a rotating panel whereby households exit the survey at the end of the fifth wave and are replaced with new households. Interviews

during the first wave are conducted face to face, and on the phone for subsequent waves. The sample covers approximately 60,000 households. Each quarter consists of approximately a fifth of the sample from each wave. The sample of the addresses are drawn from the Postcode Address File (PAF), a computer list of all delivery addresses points for Royal Mail which receive less than 50 items of mails a day – addresses which are later identified to be non-residential are excluded from the sample. PAF is sorted by postcode, i.e. geographically. The ONS suggest that PAF covers 97 per cent of private households in Britain, and about 1.5% of the total population in Britain may not be accounted for in the LFS because of not sampling from communal establishments, which tend to have higher proportions of women, older people and lower employment rates. The survey may not also provide a good representation of recent immigrants because of its sample size (Clark and Drinkwater, 2008).

The attractiveness of the LFS is that it contains a larger sample of immigrants than any other surveys, and as such it has been widely used in studies of labour market outcomes of immigrants (Longhi and Rokicka, 2012; Algan et al, 2010; Drinkwater et al, 2009; and Clark and Drinkwater, 2008). Because the survey is a rotating panel which follows the same individual over five consecutive quarters, we only keep the observation from the quarter when the individual is observed in the survey for the first time in order to avoid double counting. In addition to information on the country of birth of respondents, the LFS also contains the year foreign-born individuals entered the UK. We use this information to construct the age at which an immigrant entered the UK and to obtain their length of stay in the country. We subsequently only keep immigrants who entered the UK at age 21+ and who are not in full-time education. These are individuals likely to have entered the UK after completion of their full-time education, and hence more likely to have immigrated for economic reasons. To maintain comparability, the natives in the data are restricted to individuals of age 21+ and not in full-time education.

The LFS provides various measures of educational attainments such as highest qualification obtained and age left continuous full-time education. However, because a large proportion of immigrants have their qualifications reported as “other”, the age at which the respondent left continuous full-time education is

generally used as measure of educational attainment (see Algan et al (2010) and Drinkwater et al (2009)). To reduce the likelihood of misreporting, we further exclude from the sample the few individuals who report having left continuous full-time education at an age greater than 34.

Using information on country of birth, we group immigrants into three categories: EU14 (pre-existing members of the EU before 2004, which excludes non EU15 members of the European Economic Area as visa requirements vary across this group); A8 (Eastern and Central European countries that obtained unrestricted access to the UK labour market from May 2004 onward); and “non-EU” immigrants, which refers to immigrants from outside the EEA zone. The data we use span the period of 2001 to 2009. The choice of this time period is motivated by the fact that we want to compare the occupational distributions of immigrants who entered the UK shortly before the 2004 change in the immigration policy to that of immigrants who entered the UK shortly after the policy change.

The LFS provides occupation codes at the 3-digits, 2-digits and 1-digit levels. The Office of National Statistics (ONS), which conducts the survey, also provides 4 categories of occupations based on the skill requirements of the job<sup>9</sup>. Table 2.2 provides a detailed description of what occupations belong to which categories, with occupations at the 2 digit standard occupational classification (SOC2000). We see that skill level 4 consists of professional and managerial occupations. Skill level 3 comprises associate professional and skill trade occupations. Skill level 2 is made up of sale, personal services and machine operation occupations. Skill level 1 consists of elementary occupations. In this chapter, we refer to skill level 4 occupations as “professional” occupations, skill level 3 occupations as “semi-professional” occupations, skill level 2 occupations as “skilled manual/routine” occupations, and skill level 1 occupations as “elementary” occupations.

Table 2.3 shows the distributions of A8, EU14 and non-EU immigrants across these four occupation categories using LFS data covering the period of 2001 to 2009. The table shows that the different groups of immigrants do not have the same occupational distributions, either before or after 2004. Immigrants from

---

<sup>9</sup> See page 3 of the “Standard Occupational Classification 2010, Volume 1, Structure and descriptions of unit groups”: <http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/index.html>

EU14 are more concentrated in professional occupations, whereas immigrants from non-EU countries are equally spread across professional, semi-professional and skilled manual/routine occupations. As for A8 immigrants, they are mostly concentrated in semi-professional, skilled manual/routine and elementary occupations.

Table 2.3 also contrasts the occupational distributions of those who entered the UK before and after 2004 for each group of immigrants. We see that there was a sharp increase in the share of A8 immigrants in elementary occupations, rising from 24.5 per cent to over 41.2 per cent. On the other hand, the share of those employed in semi-professional occupations dropped from 30 per cent to 19 per cent. There was also a shift in the share of those employed in professional occupations – it went from almost 10 per cent to 4.1 per cent. The shares of those employed in skilled manual/routine occupations remained almost the same. Turning to immigrants from non-EU countries, we see that the difference in occupational distributions of those who entered the UK before and after 2004 is less pronounced. The most noticeable difference is that amongst those who entered the UK after 2004, the share of employment in elementary occupations was 18.3 per cent compared to 13.3 per cent for their counterparts who entered the UK before 2004. Semi-professional occupations accounted for about 24 per cent of employment amongst those who entered the UK after 2004 compared to 28.3 per cent for their counterparts who entered the UK before 2004. For immigrants from EU14 countries, the most noticeable difference is that amongst those who entered the UK after 2004 too; the share of those employed in elementary occupations was 14.2 per cent compared to just below 11.1 per cent for their counterparts who entered the UK before 2004. And the share of those employed in a skilled manual/routine occupation was about 20.5 per cent amongst those who entered the UK after 2004 compared to 27 per cent for their counterparts who entered the UK before 2004.

What is more surprising, however, is that the age at arrival and the age of completion of continuous full-time education remained relatively constant within each group over the years. Table 2.4 shows the median values for these measures of human capital for each yearly cohorts that entered the UK between 2001 and 2008. The median age at arrival was between 25 and 28 amongst A8 immigrants, between 28 and 30 amongst non-EU immigrants and between 27 and 29

amongst immigrants from EU14 countries across these 8 years. Median age of completion of continuous full-time education at arrival was consistently around 20 amongst A8 immigrants, 21 amongst non-EU immigrants and 21 (or 22) amongst immigrants from EU14 countries across these 8 years.

**Table 2.2: Classification of occupations based on skill requirements for task performed as per the ONS**

| SOC 2000 2-digit occupations                             | Skill level category |
|----------------------------------------------------------|----------------------|
| Corporate managers                                       | 4                    |
| Science and technology professional                      | 4                    |
| Health professionals                                     | 4                    |
| Teaching and research professionals                      | 4                    |
| Business and public service professionals                | 4                    |
| Managers and proprietors in agriculture and services     | 3                    |
| Science and technology associate professional            | 3                    |
| Health and welfare associate professionals               | 3                    |
| Protective services occupations                          | 3                    |
| Culture, media and sport occupations                     | 3                    |
| Business and public service associate professionals      | 3                    |
| Skilled agriculture trades                               | 3                    |
| Skilled metal and electrical trades                      | 3                    |
| Skilled construction and building trades                 | 3                    |
| Textiles, printing and other skilled trades              | 3                    |
| Administrative occupations                               | 2                    |
| Secretarial and related occupations                      | 2                    |
| Caring personal service occupations                      | 2                    |
| Leisure and other personal service occupations           | 2                    |
| Sales occupations                                        | 2                    |
| Customers service occupations                            | 2                    |
| Process, plant and machine operative occupations         | 2                    |
| Transport and mobile machine drivers and operatives      | 2                    |
| Elementary trades, plant and storage related occupations | 1                    |
| Elementary administrative and service occupations        | 1                    |

*Based on Standard Occupational Classification 2010, Volume 1, Structure and descriptions of unit groups.*

**Table 2.3: Occupational distributions by time of arrival and place of origin – LFS (2001-2009)**

| <i>Immigrants from</i>             | Before 2004 arrivals |        |       | After 2004 arrivals |        |       |
|------------------------------------|----------------------|--------|-------|---------------------|--------|-------|
|                                    | A8                   | Non-EU | EU14  | A8                  | Non-EU | EU14  |
| Professional occupations           | 9.97                 | 28.02  | 36.91 | 4.1                 | 28.42  | 39.33 |
| Semi-professional occupations      | 30.01                | 28.26  | 25    | 19.15               | 23.96  | 25.97 |
| Skilled manual/routine occupations | 35.5                 | 30.42  | 26.98 | 35.52               | 29.32  | 20.47 |
| Elementary occupations             | 24.52                | 13.3   | 11.11 | 41.24               | 18.3   | 14.22 |

**Table 2.4: Median age at arrival in the UK and educational attainment levels of immigrants, by year of arrival**

| <i>Immigrants from</i> | <b>Age arrived in the UK</b> |              |                | <b>Age left continuous full-time education</b> |              |              |
|------------------------|------------------------------|--------------|----------------|------------------------------------------------|--------------|--------------|
|                        | A8                           | Non-EU       | EU14           | A8                                             | Non-EU       | EU14         |
| 2001                   | 25<br>(0.50)                 | 28<br>(0.17) | 27<br>(0.37)   | 20<br>(0.22)                                   | 21<br>(0.08) | 21<br>(0.20) |
| 2002                   | 26<br>(0.50)                 | 29<br>(0.19) | 27<br>(0.44)   | 20<br>(0.25)                                   | 21<br>(0.09) | 21<br>(0.23) |
| 2003                   | 26<br>(0.42)                 | 28<br>(0.20) | 28<br>(0.43)   | 20<br>(0.19)                                   | 21<br>(0.10) | 21<br>(0.23) |
| 2004                   | 26<br>(0.27)                 | 29<br>(0.22) | 28<br>(0.48)   | 20<br>(0.11)                                   | 21<br>(0.10) | 21<br>(0.26) |
| 2005                   | 26<br>(0.21)                 | 29<br>(0.24) | 27<br>(0.45)   | 20<br>(0.09)                                   | 21<br>(0.11) | 22<br>(0.23) |
| 2006                   | 26<br>(0.26)                 | 29<br>(0.27) | 27<br>(0.53)   | 20<br>(0.11)                                   | 21<br>(0.12) | 22<br>(0.27) |
| 2007                   | 26<br>(0.36)                 | 29<br>(0.32) | 29<br>(0.74)   | 19<br>(0.14)                                   | 21<br>(0.14) | 23<br>(0.37) |
| 2008                   | 27<br>(0.63)                 | 30<br>(0.53) | 27.5<br>(0.96) | 19<br>(0.23)                                   | 21<br>(0.23) | 22<br>(0.37) |

*Standard errors in parentheses*

## 2.6 Methodology

This section starts by presenting the choice of our estimator. We follow the formulation by McFadden (1974) in analysing discrete outcomes given our outcome variable (i.e. observed occupations of immigrants) takes a small number of discrete values. We then outline how we empirically disentangle the impact of immigration policy on the occupational outcomes of A8 immigrants from the effect that may be attributed to their place of origin. We also discuss the challenges of identifying the impact of immigration policy on our occupational outcomes of interest using cross-section data in a difference in differences framework.

### 2.6.1 Analysis of occupational outcomes

Suppose an individual  $i$  in the population has a vector of measured attributes  $\mathbf{x}_i$ , and faces  $J$  alternatives choices, indexed by  $j = 1, 2, \dots, J$ , each with corresponding vectors of attributes  $\mathbf{z}_j$ . While  $\mathbf{x}_i$  is invariant across alternatives,  $\mathbf{z}_j$  may or may not vary across individuals.

Suppose further the individual has a utility function that can be written in the form of

$$U = V(x, z) + \varepsilon(x, z) \quad (2.6.1)$$

where  $V$  is a non-stochastic component of  $U$  and  $\varepsilon$ , its random component, reflects the idiosyncrasies of the individual's taste for the alternative with attribute  $z$ . Let  $y$  denote the observed choice outcome of the individual. Assuming rational individuals who maximise utility, and denoting the probability that a given individual, chosen at random, with attribute  $x$  chooses  $j$  from the set of alternatives  $J$  as:

$$P_{ij} = P_i(y = j | x, J) = P(U_{ij} > U_{ik}), \text{ for all } j \text{ not equal to } k; \quad k \& j \in J \quad (2.6.2)$$

McFadden shows that under 1) independence of irrelevance alternatives (IIA), that is the odds of alternative  $j$  being chosen over alternative  $k$  in the presence of other alternatives in the choice set  $J$  equals the odds of a binary choice between  $j$  and  $k$ ; and 2) positivity, that is each element of  $J$  has a strictly positive probability of being chosen, and 3) that  $\varepsilon(x, z_j)$  is independently identically distributed with Extreme Type I distribution, we have:

$$\log \left[ \frac{P_{ij}}{P_{ik}} \right] = V(x_i, z_j) - V(x_i, z_k) \quad (2.6.3)$$

Empirical estimation requires specifying a functional form for  $V(x, z_j)$ , and often a linear function is assumed. With only individual attributes in the data of analysis, as is our case, this may take the form of  $V(x, z_j) = \beta_j x_i$  and it follows

$$\text{that: } P_{ij} = P_i(y = j | x, J) = \frac{e^{\beta_j x_i}}{\sum_{k \in J} e^{\beta_k x_i}} \quad (2.6.4)$$

This gives rise to the multinomial logit specification.

### 2.6.2 Disentangling the impact of immigration policy from the impact of birthplace on the occupational outcomes of immigrants

One could think of the idea underlying our approach as follows. Suppose there are two groups in the population which are observed over two periods, such that while neither of them is subject to a policy change in the first period, one of them undergoes a treatment in the second period by being subjected to a change in policy. If both groups are affected in a similar way by time (in the absence of the

policy change), then the change in the outcome of the group that underwent treatment, relative to the other group, would be attributed to the change in policy. This method of policy evaluation is called “difference in differences”. The use of difference in differences for policy/impact evaluations ranges from application to questions of health, educational policy analysis, labour market outcomes studies and many more (see Lechner (2011) for a review). As noted by Lee and Kang (2006), “difference in differences is one of the most popular – and often convincing – study designs in finding the effects of a treatment in social science.”

An application of difference in differences when the dependent variable is a discrete variable, in the context of the UK, is Blundell et al (2004). The UK government allocated employment advisors to 19-24 year olds claiming unemployment benefit and made it compulsory for them to discuss their job search progress with an advisor on a fortnightly basis. Using the policy as treatment, the paper compares the before and after employment probabilities of the treated group and those of other non-affected age groups – i.e. the control groups. They show that mandatory job search programmes increased employment probabilities by 20 per cent. Kugler et al (2002), in the context of Spain, exploit a change in payroll taxes and dismissal costs for permanent contracts to examine the impact of labour market laws on employment. By comparing the affected young workers by the change in policy to other non-affected individuals before and after the policy change, they find that employment of young worker on permanent contracts increased after the change in policy.

In our case, one may decompose the population into 4 subgroups: A, E, O and N. Subgroup A consists of immigrants from A8 countries, subgroup E consists of immigrants from EU14 countries, O consists of immigrants from other part of the world and subgroup N consists of natives. Groups E and N had unrestricted access to the UK labour market both before and after the policy change; Group O had restricted access both before and after the policy change; and Group A had restricted access before and unrestricted access after the policy change. Hence we have 3 possible control groups to examine the question of interest in a difference in differences framework. Thus we have:

$$\delta_{DiD} = (\bar{y}_{after}^A - \bar{y}_{before}^A) - (\bar{y}_{after}^O - \bar{y}_{before}^O) \quad (2.6.5)$$

where *before* denotes entry in the UK before 2004 and *after* denotes entry in the UK after 2004. In equation (2.6.5), subgroup O is used as the control group. One could also use groups N and E as control groups, since both these groups also experienced no change in policy over the time period considered. In a linear regression framework, this would be equivalent to taking a sample comprising immigrants from subgroups A and O and estimating  $\delta_{DiD}$  from:

$$y_{i,t} = \delta_0 + \delta_1 post2004 + \delta_2 A8 + \delta_{DiD} A8 * post2004 + \varepsilon_{i,t} \quad (2.6.6)$$

where *A8* is a dummy variable taking the value of 1 if the individual is from one of the A8 countries.  $\delta_{DiD}$  captures the effect of unrestricted access to the UK on the occupational outcomes of immigrants from the A8 countries following the change in policy.

### 2.6.3 Challenges of identifying the impact of immigration policy on the occupational outcome of immigrants based on independent cross section

This section discusses challenges faced by the researcher when using independent cross section data to examine the effect of a policy change on some outcome. Here, we continue with the case of a linear equation for ease of illustration.

In the case of a first difference estimation,  $\delta_{FD}$ , the analogue of  $\delta_{DiD}$ , is obtained as:

$$\delta_{FD} = \bar{y}_{after}^A - \bar{y}_{before}^A \quad (2.6.7)$$

or from

$$y_{i,t} = \alpha + \delta_{FD} Policy\_change + \varepsilon_{i,t} \quad (2.6.8)$$

$\delta_{FD}$  measures the difference in means between the outcomes before the policy came into effect and after the policy came into effect. Note that in this case, the sample would only consist of one group.

As suggest Bell et al (1999), now suppose one can decompose  $\varepsilon_{i,t}$  into two additive components: 1)  $\mu_{i,t}$ , a term capturing individual effect uncorrelated across individuals; and 2)  $k_i m_t$ , a macro time component which is allowed to

affect individuals differently. So, we have  $\varepsilon_{i,t} = \mu_{i,t} + k_i m_t$ .  $\delta_{FD}$  is consistently estimated under the assumption that

$$E[\varepsilon_{i,t}|t] = \varepsilon \text{ for all } t = \textit{after}, \textit{before}.$$

This means that the individual idiosyncratic component need not have zero mean before and after the policy change, and one can use independent cross sectional data to estimate  $\delta_{FD}$ . What is required is that 1) the sample of immigrants before the change in policy has the same composition of unobservables as the sample of immigrants affected by the change in policy, and 2) the macro condition remained the same, i.e.  $m_t = m$  (Bell et al, 1999). As noted by Bell et al (1999), the first assumption can be satisfied by carefully choosing the *before* period group according to the eligibility criteria for the policy in question and controlling for any remaining observable difference through a set of individual characteristics. The second requirement may not be satisfied if there was any systematic change in the macro economy in the *after* period, that is  $m_{\textit{after}} \neq m_{\textit{before}}$ . If this is the case, it can be addressed by using difference-in-differences.

Assume there is a control group, say another group of individuals, either natives or immigrants, who were not affected by the change in immigration policy. In our setting, we can have this group as O – i.e. non-EU immigrants – and we have immigrants from A8 countries, the treatment group as A. The impact of the policy can then be estimated by differencing out (2.6.7) for group A and O, giving us the difference in differences estimate ( $\delta_{DiD}$ ) in (2.6.5). Consistent estimation of  $\delta_{DiD}$  requires that:

$$E\{\varepsilon_{i,t}|t, i \in g\} = u_g + k m_t, \text{ for all } g \text{ and } t.$$

$g$  denotes the group an individual belong to, i.e. the control group or treatment group.

Two conditions are required: 1) in the absence of the treatment, both the control and treatment would have followed similar trends, and 2) there is no compositional change amongst the group treated such that the unobserved group effect,  $u_g$ , is the same in the *before* and *after* period. In other words,  $u_g$  does not change *before* and *after* the treatment. We will examine the validity of these two

conditions later in the essay – initial observations from section 2.4 showed that the median age and median educational attainment levels at arrival in the UK, measures of human capital, appear not to have changed much in the *before* and *after* periods.

#### 2.6.4 The impact of policy on occupational outcomes

As noted earlier, our outcome variable takes a small number of finite discrete values and our estimator is non-linear. An underlying assumption using difference in differences estimation is that the error term can be separated into a time component and a group component. The error term is then assumed to be additive and separable in these two terms. The linearity assumption between the outcome variable and the error term may not necessary hold in non-linear cases. A common practice, as noted by Puhani (2012), is to apply the assumptions made in linear cases to the latent index of the outcome variable rather than the observed outcome itself. The study of the impact of mandatory job search programme on employment probability by Blundell et al (2004) is an example of such an approach. They impose the linearity assumption on the latent index function to estimate the coefficient of interest using difference in differences.

The approach we take in examining the impact of the immigration policy on the occupational outcomes of immigrants is to compare the occupational distributions of A8 immigrants before and after the policy change, relative to those of other immigrants that were not affected by the change in policy. We can do this by looking at how odds ratios of being in an occupation, relative to a fixed chosen occupation, changed. The log transformation of equation (2.6.4) leads to:

$$\log[P_j/P_k] = X(\beta_j - \beta_k) \quad (2.6.9)$$

Choosing a base alternative, say  $k$ , such that  $\beta_k = 0$ , we have

$$\log[P_j/P_k] = X\beta_j \quad \text{for } j \neq k \quad (2.6.10)$$

Equation (2.6.10) is a linear equation similarly to equation (2.6.6). This allows us to examine the impact of freedom from immigration restriction on the

occupational outcomes of immigrants. More specifically, in our context, the basic of form of this will be:

$$\log[P_j/P_k] = \beta_{j0} + \beta_{j1}\text{post2004} + \beta_{j2}A8 + \beta_{\text{DiD}}\text{post2004} * A8 \quad (2.6.11)$$

Another important consideration by the researcher is the suitability of the control group (Angrist and Pischke, 2009). The ideal control group is one that would arguably respond to the treatment in a very similar way to the treated group. Of all the three potential control groups presented earlier (natives, EU14 immigrants and non-EU immigrants), the non-EU group of immigrants has the advantage of still requiring some form of leave to enter and/or work in the UK. We use this group as our main control group and examine the robustness of our results using immigrants from EU14 countries.

## 2.7 Empirical analysis and results

This section starts by examining the assumptions necessary for equation (2.6.5) to consistently estimate the effect of the policy change. We then proceed to examining the effect of the change of policy on the occupational outcomes of A8 immigrants. In order to make the before and after groups more comparable, we restrict the sample to immigrants who entered the UK between 2001 and 2006 (i.e. shortly before and after the change in policy) and only keep those with a duration of stay in the UK of no more than 5 years (at the time of the survey) in order to strike a balance between having ample sample size and comparing immigrants with similar duration of stay in the UK. Summary statistics of our sample are presented in Table 2.5.

We see from the table that A8 immigrants are on average younger than EU14 immigrants, who are in turn on average younger than immigrants from non-EU countries. This remains the case both before and after 2004. The same pattern is observed for age at arrival. The means of age of completion of continuous full-time education are largely comparable across the three groups of immigrants, before and after 2004.

Unlike Table 2.5 where we restrict the sample to immigrants who have lived in the UK for at most 5 years, Table 2.3 does not account for duration of stay.

When one contrasts occupational distributions from Table 2.5 with those from Table 2.3, we see that the share of A8 immigrants who entered the UK before 2004 and were employed in elementary occupations is 24.5 per cent in Table 2.3 as opposed to 31.1 per cent in Table 2.5, a difference of almost 7 percentage points. Similarly, the share of non-EU immigrants in elementary occupations increases from 13.3 per cent in Table 2.3 to 16.2 per cent in Table 2.5. The corresponding increase amongst immigrants from EU14 countries is over 8 percentage points, i.e. 11.1 per cent in Table 2.3 versus 19.3 per cent in Table 2.5. Turning to professional occupations, we see that the shares of those employed in these occupations are greater in Table 2.3 than in Table 2.5 across all groups of immigrants who entered the UK before 2004.

The occupational distributions in Table 2.5 and Table 2.3 are quite comparable for immigrants who entered the UK after 2004. This is partly because in both tables all immigrants would have lived in the UK for at most 5 years, but the reported figures in Table 2.3 for those who entered the UK before 2004 comprise immigrants who have lived in the UK for over 5 years, unlike those in Table 2.5. These contrasts in occupational distributions suggest that duration of stay in the host country potentially plays a role in the occupational outcomes of the immigrants considered here, and therefore any comparison of occupational outcomes before and after 2004 should take into account length of stay in the UK.

To further illustrate the importance of accounting for length of stay in the UK in our analysis, Figure 2.1 presents the occupational distribution of all A8 immigrants who entered the UK from 1995 onward and started a spell of employment after May 2004, the date they became free from immigration restrictions in the UK. These occupational distributions are compared across 4 groups, differentiated by the length of stay of the immigrant in the UK: 0 to 2 years, 3 to 5 years, 6 to 10 years, and over 10 years.

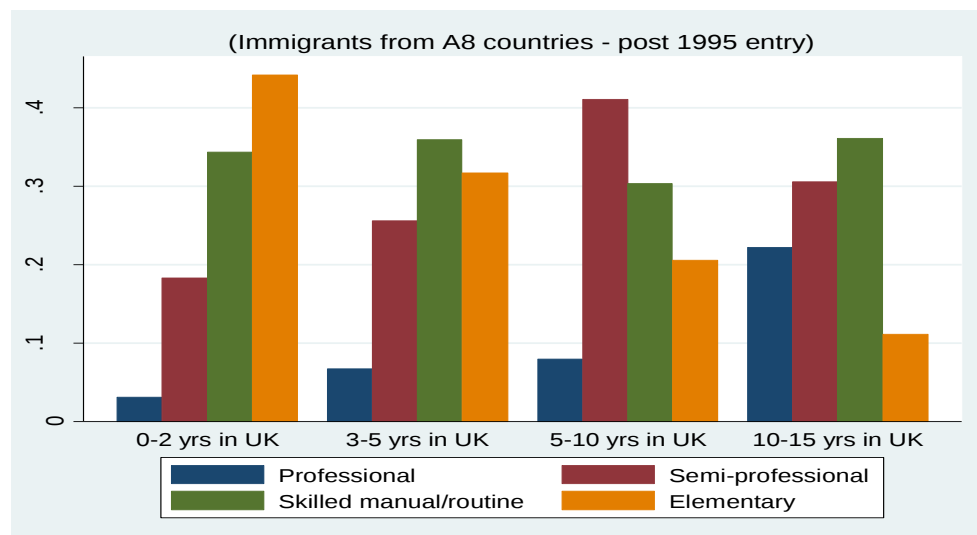
We see a downward trend in the share of immigrants employed in an elementary occupation with duration of stay in the UK, but the opposite in the share of those in a professional occupation and to a large extent in the share of those in a semi-professional occupation. The share of those in a skilled manual/routine occupation appears fairly stable over time. Amongst immigrants with over 10

years of stay in the UK, elementary occupations accounts for only 10 per cent compared to about 20 per cent amongst immigrants with 5 to 10 years of stay in the UK and over 40 per cent amongst immigrants with 0 to 2 years of stay in the UK. We also see that amongst immigrants who have been in the UK for 0 to 2 years, close to 35 per cent are in a skilled manual/routine occupation, over 15 per cent are in a semi-professional occupation and less than 5 per cent are in a professional occupation. On the other hand, looking at immigrants who have been in the UK for 5 to 10 years, about 10 per cent are in a professional occupation. This rises to about 20 per cent amongst immigrants who have been in the UK for over 10 years.

The patterns observed in Figure 2.1 clearly show that immigrants with longer duration of stay in the UK have better occupational outcomes, and therefore any comparison of occupational outcomes before and after 2004 should take this into account. In short, the very large change in occupational distribution before and after 2004 shown in Table 2.3 is at least in part related to the fact that the immigrants before 2004 had been in the UK for longer.

**Table 2.5: Summary statistic of the sample: Immigrants who entered the UK between 2001 and 2006, LFS (2001-2009).**

|                                                    | A8<br>immigrants | Non-EU<br>immigrants | EU14<br>immigrants |
|----------------------------------------------------|------------------|----------------------|--------------------|
| <b>Entered the UK before 2004 (i.e. 2001-2003)</b> |                  |                      |                    |
| <b>Individual characteristics</b>                  |                  |                      |                    |
| Mean age of arrival                                | 27.68<br>(0.31)  | 30.58<br>(0.12)      | 29.88<br>(0.26)    |
| Mean age                                           | 30.54<br>(0.32)  | 33.17<br>(0.12)      | 32.10<br>(0.27)    |
| Age left cont. full-time education                 | 20.36<br>(0.14)  | 20.59<br>(0.06)      | 20.88<br>(0.14)    |
| <b>N</b>                                           | 467              | 4,062                | 887                |
| <b>Occupational distribution (%)</b>               |                  |                      |                    |
| Managerial and professional                        | 7.51             | 27.64                | 32.96              |
| Assoc. professional & skill trade                  | 27.47            | 28.57                | 23.14              |
| Semi-skill                                         | 33.91            | 27.51                | 24.60              |
| Elementary                                         | 31.12            | 16.28                | 19.30              |
| <b>Entered the UK after 2004 (i.e. 2004-2006)</b>  |                  |                      |                    |
| <b>Individual characteristics</b>                  |                  |                      |                    |
| Mean age of arrival                                | 28.58<br>(0.14)  | 31.17<br>(0.14)      | 29.96<br>(0.28)    |
| Mean age                                           | 30.63<br>(0.14)  | 33.34<br>(0.14)      | 31.91<br>(0.28)    |
| Age left cont. full-time education                 | 20.09<br>(0.06)  | 20.76<br>(0.06)      | 21.61<br>(0.15)    |
| <b>N</b>                                           | 2,813            | 3,215                | 715                |
| <b>Occupational distribution (%)</b>               |                  |                      |                    |
| Managerial and professional                        | 4.40             | 27.59                | 39.30              |
| Assoc. professional & skill trade                  | 19.52            | 25.26                | 25.73              |
| Semi-skill                                         | 36.23            | 29.6                 | 20.00              |
| Elementary                                         | 39.85            | 17.54                | 14.97              |

*Standard errors in parentheses***Figure 2.1: Occupational distribution of A8 immigrants by duration of stay in the UK**

## 2.7.1 Plausibility of the underlying of assumptions in section 2.6.3

### 2.7.1.1 The effect of policy on the composition of immigrants

This section examines whether the composition of A8 immigrants systematically changed following the change in the UK's immigration policy. We focus on two characteristics that we are able to observe in the data: age at arrival in the UK and educational attainment levels. Table 2.4 shows that the median age at arrival in the UK and the educational attainment levels of immigrants from A8 countries remained relatively unchanged from 2001 to 2009, as did that of other groups. Additionally, looking at the mean values across the before and after 2004 groups of immigrants in Table 2.5, we see very little changes. Immigrants from A8 countries who entered the UK after 2004 are, on average, 0.9 years older than immigrants who entered the UK before 2004; the difference in means of age left continuous full-time education is 0.3. Changes in characteristics did not only occur amongst A8 immigrants, however. Table 2.5 shows that immigrants from non-EU countries who entered the UK after 2004 were, on average, 0.6 years older and had 0.2 more years of continuous full-time education than their counterparts who entered the UK before 2004. Among EU14 immigrants, the difference in means of age at arrival in the UK between the before and after 2004 groups is 0.1 years and the difference in means of educational attainment levels is 0.8.

Table 2.6 presents results from a formal test of differences in the mean values reported in Table 2.5. We do so by estimating the following regression:

$$y_{i,t} = \alpha + \partial_1 \text{post2004} + \varepsilon_{i,t} \quad (2.7.1)$$

$y_{i,t}$  is either the age at arrival in the UK or the educational attainment level of the individual. As before, *post2004* is a dummy variable which takes the value 1 if the immigrant entered the UK following the policy change. The equation is estimated for each group of immigrants separately – recall that the sample comprises immigrants who have lived in the UK for at most five years at the time of the survey, and only those who entered the UK between 2001 and 2006 are kept in the sample.  $\partial_1$  corresponds to differences in mean values we have just discussed above (i.e. those reported in Table 2.5). The advantage of estimating equation

(2.7.1) is that it is a straight forward way of gauging the statistical significance of these differences.

None of the differences comes out statistically significant at the conventional level for A8 immigrants, so we cannot reject that the characteristics looked at remained largely the same before and after 2004 for this group. For non-EU immigrants, the difference in means of age at arrival in the UK is statistically significant, albeit as seen earlier, the magnitude is very small. For their mean age of leaving continuous full-time, the difference between the before and after 2004 values is not statistically significant. Turning to immigrants from EU14 countries, we see the reverse pattern: the difference in means of age at arrival in the UK is not statistically significant, but for their mean age of leaving continuous full-time, the difference in the before and after 2004 values is statistically significant.

We can compute from Table 2.6 by how much changes in the characteristics of A8 immigrants differ to changes in characteristics amongst other groups of immigrants. The values for these and their statistical significance levels are reported in Table 2.7 for ease of exposition. The relative increase in age at arrival amongst A8 immigrants is 0.31 (0.82) when compared with non-EU immigrants (EU14 immigrants). In addition to the small magnitude of these coefficients, they are statistically insignificant, suggesting lack of evidence of systematic change in the age at arrival of A8 immigrant following the change in policy. When we compare means of age left continuous full-time education, the change amongst A8 immigrants is 0.44 lower than the change amongst immigrants from non-EU14 countries. When compared to EU14 immigrants, the corresponding value is 1.07 lower. With a mean age of completion of full-time education of over 20 years, these changes are relatively quite small. As such, here too we see very little to no systematic change in the educational attainment levels of A8 immigrants who entered the UK after the change in immigration policy.

For further additional checks, we repeat the exercise conducted above by 1) omitting immigrants who entered the UK in 2004 from the sample given the change in policy occurred in May 2004 and we do not observe the month when an immigrant entered the UK from the data; and 2) looking at the stock of immigrants in the UK between 2001 and 2006. We reach similar conclusion:

there was no systematic change in the characteristics of A8 immigrants following the change in policy. See Tables 2.14 to 2.17 in the appendix for detail.

We are aware that the selection of immigrants through visa policy may take into account additional observable characteristics to the policy maker such as (English) language proficiency, experience in particular fields or trade, income levels, etc. We are unable to observe these characteristics in the data at hand, hence why we only look at educational attainment levels and age at arrival in the UK – language proficiency is observed for only a small number of observations in our data and this does not allow meaningful analysis. We are mindful that if these other characteristics systematically differ amongst immigrants who entered the UK before and after the change in visa policy, and are positively/negatively associated with their occupational outcomes in the UK, this may have implications for our analysis. Nevertheless, one would expect these characteristics to be highly correlated with educational attainment levels and the age of individuals at entry in the UK, which would negate the extent of the resulting biases, or at least reduce them.

**Table 2.6: Difference in means of characteristics before and after 2004, by immigrant group. Only immigrants who entered the UK between 2001 and 2006**

|                    | Age at arrival in the UK,<br>immigrants from |                       |                       | Age left-continuous full time<br>education, immigrants from |                       |                       |
|--------------------|----------------------------------------------|-----------------------|-----------------------|-------------------------------------------------------------|-----------------------|-----------------------|
|                    | A8                                           | Non-EU                | EU14                  | A8                                                          | Non-EU                | EU14                  |
| After 2004 arrival | 0.896<br>(1.620)                             | 0.585**<br>(2.098)    | 0.072<br>(0.166)      | -0.275<br>(-1.338)                                          | 0.168<br>(1.391)      | 0.732***<br>(4.809)   |
| Mean               | 27.683***<br>(41.827)                        | 30.581***<br>(82.948) | 29.888***<br>(33.836) | 20.362***<br>(45.420)                                       | 20.595***<br>(94.919) | 20.874***<br>(27.642) |
| N                  | 3,280                                        | 7,277                 | 1,602                 | 3,280                                                       | 7,277                 | 1,602                 |

Statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . *t*-statistics in bracket. Standard errors clustered by country of birth. Sample comprises those who have lived in the UK for at most 5 years at the time of the survey.

**Table 2.7: Test of differences in changes across groups reported in Table 2.6**

|           | Age at arrival in the UK |                  | Age left-continuous full time education |                       |
|-----------|--------------------------|------------------|-----------------------------------------|-----------------------|
|           | A8 vs non-EU             | A8 vs EU14       | A8 vs non-EU                            | A8 vs EU14            |
| Magnitude | 0.311<br>(0.528)         | 0.824<br>(1.208) | -0.443*<br>(-1.946)                     | -1.007***<br>(-4.064) |

Statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . *t*-statistics in bracket. Standard errors clustered by country of birth. Sample comprises those who have lived in the UK for at most 5 years at the time of the survey.

### 2.7.1.2 The validity of commonality of time trend prior to the policy change

This section examines whether or not the trend in occupational distributions of A8 immigrants who entered the UK prior to the change in policy was similar to those of other immigrants. To this purpose, we look at the trends in the occupational distributions of immigrants who entered the UK three years before the change in immigration policy. We also present trends in these distributions for those who entered the country three years after the change in policy to provide an illustrative picture on whether these remained similar to trends for the before 2004 arrivals. Here too, our sample consists of immigrants who have lived in the UK for at most 5 years at the time of the survey to strike a balance between sample size and comparability of immigrants with similar duration of stay in the UK. Figure 2.2 compares immigrants from A8 countries and immigrants from non-EU countries, and Figure 2.3 compares immigrants from A8 countries and immigrants from EU14 countries. Both figures contains four graphs, each pertaining to one of the four occupations we look at. The x-axis represents cohorts of immigrants by time of arrival in the UK, and the y-axis reports the share of employed immigrants in a particular occupation.

Figure 2.2 shows that the trend in occupational distributions of A8 immigrants was quite similar to those of non-EU immigrants for cohorts who entered the UK up until cohorts of 2003, the year preceding the change in the UK's immigration policy, before showing minor differences: the shares of those employed in a professional occupation or an elementary occupation show similar patterns over the yearly cohorts, with differences in magnitudes of employment shares up until 2003 as suggests the 95% confidence intervals. The shares of those employed in a skilled manual/routine occupation also exhibit similar patterns across the yearly cohorts of arrivals, but the 95% confidence intervals of the magnitudes of employment shares overlap until cohorts of 2003. As for shares of employment in semi-professional occupations, we observe differences across groups, but the continual large overlaps in 95% confidence intervals of the magnitudes of employment shares in these occupations across groups entail that we cannot rule out similarity in trends.

Figure 2.2: Trend of occupational distributions over time, by group of immigrants: A8 and non-EU immigrants (vertical lines indicate 95% CI range)

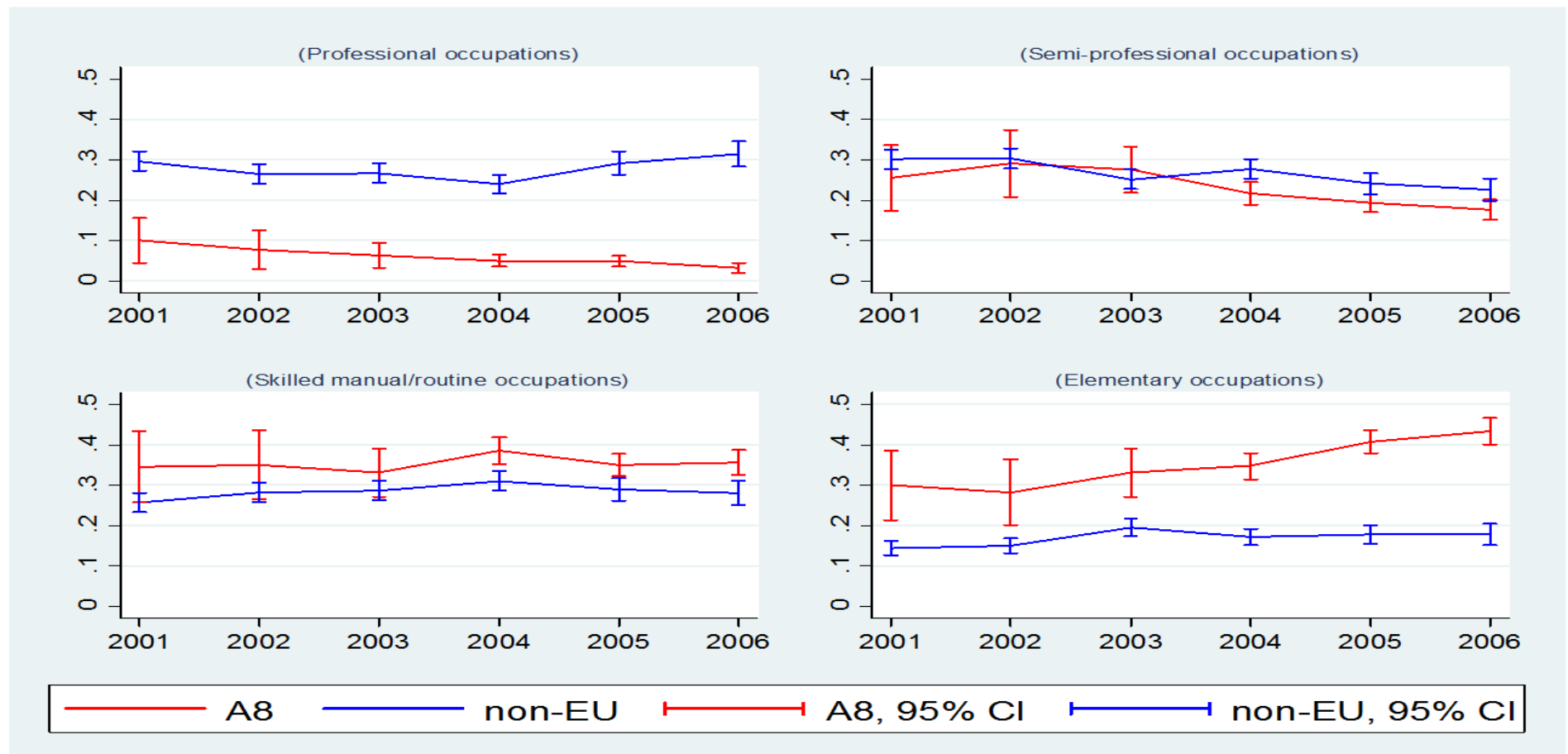
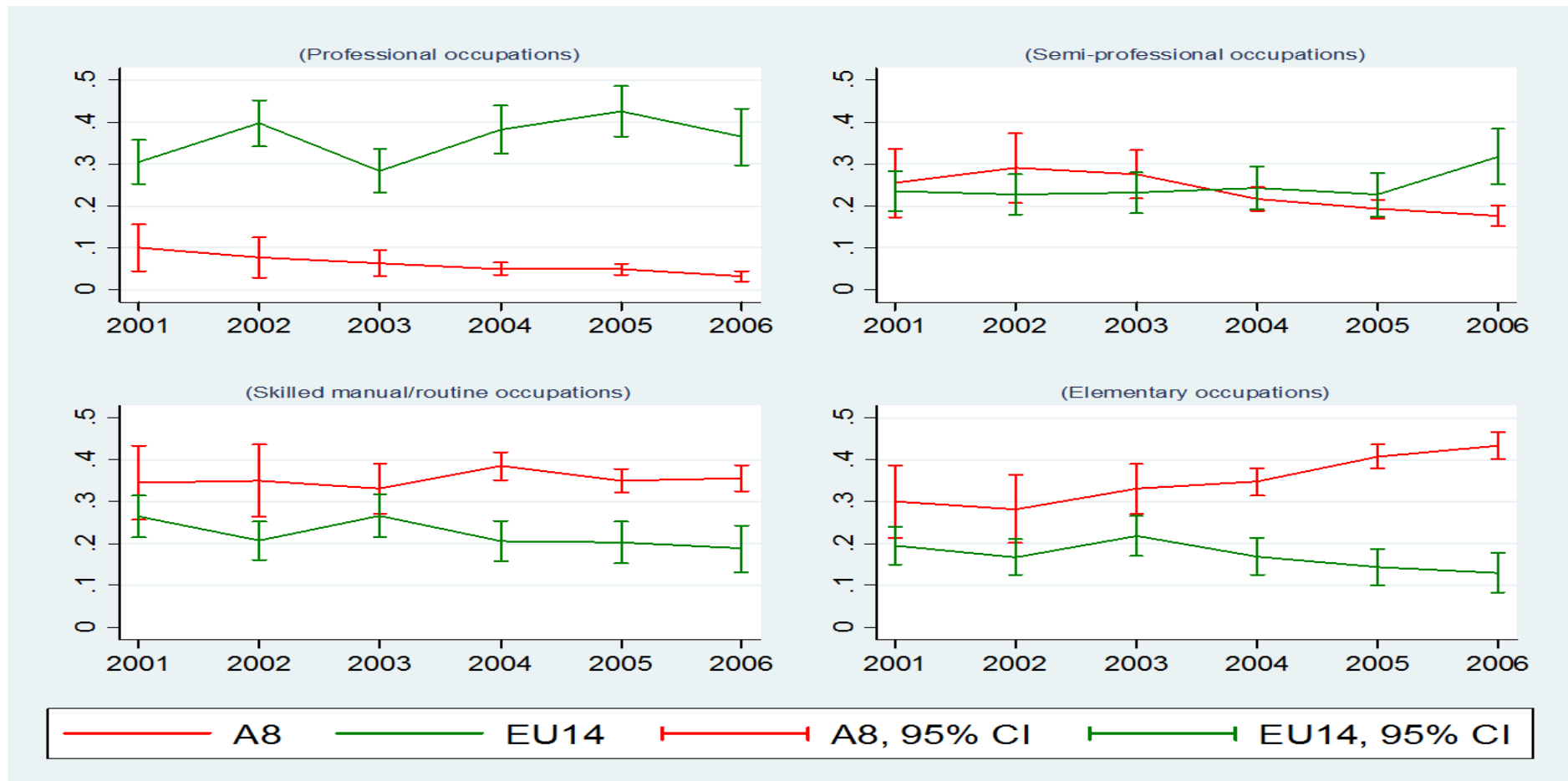


Figure 2.3: Trend of occupational distributions over time, by group of immigrants: A8 and non-EU immigrants (vertical lines indicate 95% CI range)



As a whole, we conclude that the patterns observed in Figure 2.2 suggest that A8 immigrants largely exhibited similar trends in occupational distributions to non-EU immigrants for cohorts who entered the UK in the three years preceding the change in the UK's immigration policy.

Turning to Figure 2.3, we observe clear similarity in the trend of employment shares in elementary occupations for A8 and EU14 immigrants up until cohorts of arrivals in 2003. Employment shares in professional occupations and skilled manual/routine occupations exhibit fluctuations amongst EU14 and show different trends in comparison to A8 immigrants. However, if one were to draw straight lines in employment shares in these occupations between cohorts of arrivals in 2001 and 2003, amongst EU14 immigrants, these would be parallel to the corresponding lines for A8 immigrants. As in Figure 2.2, we observe differences for shares of employment in semi-professional occupations across the two groups, but with continual large overlaps in 95% confidence intervals of the magnitudes of employment shares in these occupations. So here too, we cannot rule out similarity in trends between arrival cohorts of 2001 to cohorts of 2003, but it emerges that immigrants from non-EU14 countries would be a more appropriate control group.

Contrasting Figure 2.2 and Figure 2.3, we see that: 1) trends in the shares of employments in professional occupations and elementary occupations amongst A8 immigrants started deviated from the corresponding trends amongst non-EU14 and EU14 immigrants from 2004 arrivals; and 2) trends in occupational distributions were similar across non-EU14 and EU14 immigrants for 2004 and 2005 arrivals. It is, however, apparent from these two figures that non-EU immigrants would be a more suitable counterfactual group in comparison to EU immigrants for our purpose because of the similarity in trends of occupational distributions amongst the before 2004 arrivals.

### **2.7.2 Immigration policy and the occupational outcomes of immigrants**

This section starts with a baseline model, which is equivalent to manually computing the effect of the policy change on the occupational outcomes of A8 immigrants, using the probabilities of observing individuals in the treated and

control groups before and after the change in policy, in different occupations. The only difference is that one has the possibility to control for given factors. We re-iterate that our preferred control group is immigrants from non-EU countries because 1) this group of immigrants still face the general immigration constraints that were previously faced by A8 immigrants before the change in policy, 2) the relative change in the composition of immigrants from A8 countries is consistently negligible and insignificant wherever this group is used as a comparison group, and 3) the pre-treatment time trend in occupational distributions is mostly the same for A8 immigrants as for immigrants from non-EU countries. Nevertheless, we also proceed further by using immigrants from EU14 countries as control group to generally examine the robustness of our results to the identity of the control group.

We carry out additional robustness checks by refining the post-treatment period. Given that immigrants from A8 countries only had freedom of access to the UK labour market from May 2004 and we are unable to distinguish in the data those who entered the UK between January and April 2004 from those who entered the UK between May and December 2004, we drop all individuals who entered the UK in 2004 for robustness check. Although most 2004 arrivals would have occurred after April 2004 amongst A8 immigrants, this is to avoid any potential classification error. Further additional robustness checks are performed using alternative estimators such as an ordered logit and a linear probability model. Throughout, we extend the baseline model to include individual characteristics in order to control for the small differences in characteristics observed in Table 2.6 and Table 2.7. As noted by Angrist and Pischke (2009), the inclusion of individual characteristics in our estimation has the advantage of increased precision of the estimate of interest. It also allows one to adjust for observable differences between the observations in different groups, and may reduce the residual variance of the coefficient of interest (Meyer, 1995).

As seen in section 2.3, empirical evidence attests that immigrants move up the occupation ladder with time spent in the host country. In order to avoid any implication of duration of stay in the host country on labour market outcomes, we further refine the sample such that it consists only of immigrants who would have had lived in the UK for no more than 2 years at the time of the survey, as

opposed to 5 years. By doing so, we are comparing new immigrants who are still accustoming themselves with the UK labour, the only distinction being that some entered the UK before 2004 and other entered after 2004.

### 2.7.2.1 Immigration policy and the occupational outcomes of immigrants: baseline model

The baseline line equation we estimate is:

$$\log[P_j/P_k] = \beta_{j0} + \beta_{j1}\text{post2004} + \beta_{j2}\text{A8} + \beta_{\text{DiD}}\text{post2004} * \text{A8} \quad (2.7.4)$$

$\beta_{j0}$  is a vector of constants, and the remaining  $\beta_j$ s measure the relationship between the dependent variables and the probability of choosing the  $j^{\text{th}}$  outcome relative to the base outcome,  $k$ . Our coefficients of interest are the vector  $\beta_{\text{DiD}}$ . We cluster standard errors by country of birth.

The second column of Table 2.8 reports the exponential values of  $\beta_{\text{DiD}}^{10}$ , i.e. the direct effect of the policy change on  $P_j/P_k$ . We have chosen professional occupations as the base occupation. Therefore, these estimates represent the odds ratios of being in a given occupation relative to professional occupations. We see from Table 2.8 that the odds of observing A8 immigrants in a semi-professional occupation relative to the base occupation did not change much following the change in immigration policy. This appears to also be the case for skilled manual/routine occupations; although the odds of A8 immigrants being observed in this occupation relative to a professional occupation was 1.3, the estimate lacks statistical significant at the 10 per cent level. On the other hand, the odds ratio of observing A8 immigrants in an elementary was 1.55. This change is a relatively small increase considering that A8 immigrants are generally about 8 times more likely to be in elementary occupations relative to professional occupations, when compared with the control group.

---

<sup>10</sup> The negative t-statistics in parentheses in the table result from reporting odd ratios, hence the negative sign for values less than 1 – i.e.  $\beta_{\text{DiD}}$  less than zero.

### 2.7.2.2 Immigration policy and the occupational outcomes of immigrants: the role of duration of stay and other characteristics

We saw earlier that the composition of A8 immigrants did not systematically change relative to other groups of immigrants. Estimates from the baseline should therefore remain robust to the inclusion of individual characteristics. We extend equation (2.6.4) to include individual characteristics and estimate the following equation:

$$\log \left[ \frac{P_j}{P_k} \right] = \beta_{j0} + \beta_{j1} \text{post2004} + \beta_{j2} \text{A8} + \beta_{\text{DiD}} \text{post2004} * \text{A8} + \partial_{j1} \text{age} + \partial_{j2} \text{age\_arrived} + \partial_{j3} \text{education} + \partial_{j4} \text{sex} \quad (2.7.5)$$

Estimates of the exponential values of  $\beta_{\text{DiD}}$  are reported in column (2) of Table 2.8. Here too, standard errors are clustered by country of birth. We see from (2) that there is very little change in the initial estimates obtained from the base model. This to a large extent corroborates the findings in section 2.7.1, where we examine the no compositional change assumption. The main noticeable change is that estimates in column (2) are now closer to 1 (i.e. the underlying coefficients are closer to 0) in comparison to those obtained from the baseline.

As a first step robustness check, we conduct a placebo test using immigrants who entered the UK between 1998 and 2003. Here too, the sample is restricted to those who have been in the UK for at most 5 years. Individuals are grouped into two groups according to their year of arrival to the UK: 1998-2000 arrivals and 2001-2003 arrivals. Our fake annulment of visa restrictions is assumed to have occurred in 2001. Table 2.18, in the appendix, reports our coefficients of interest. We see from the table that the coefficients are all very close to 1 (and statistically insignificant), suggesting that the estimated effect of the “fake” change in policy is somewhat comparable to those reported in Table 2.8.

We want to explicitly account for the potential that duration of stay in the host helps immigrants acquire host country specific human capital and move up the occupation ladder, as found in other contexts (i.e. Green et al (1999) for Canada, Chiswick et al (2005) for Australia, and Amuedo-Dorentes and De la Rica (2007) for Spain). Moreover, while Figure 2.1 shows a clear positive association between duration of stay and better occupational attainment levels, the mean duration of stay for the before 2004 arrivals in the sample (i.e. 3.72) is almost double that of

the after 2004 arrivals (i.e. 2.05). We therefore restrict our sample to immigrants with 0 to 2 years of stay in the UK, hence avoiding the potential of our estimates being upward biased by comparing immigrants with shorter length of stay in the UK to immigrants with longer length of stay. Chiswick (1978) suggest that immigrants may take up to 15 years to reach the same level of earnings as the natives. In light of this, considering immigrants with 0 to 2 years of stay in the UK to still be accustoming themselves with the UK's labour market seems reasonable. We re-estimate the baseline model with this restricted sample as well as equation (2.7.5). Coefficient estimates from these equations are reported in columns (3) and (4) of Table 2.8.

First, we see that the estimates of odds ratios from equations (2.7.4) and (2.7.5) are very similar in magnitude for any given occupation. Second, they are smaller in magnitude compared to when the sample included all immigrants with at most 5 years of stay in the UK. Third, these estimates of the coefficients of interest are all approximately 1 and insignificant, implying that the change in policy had virtually no effect on the occupational outcomes of immigrants from the A8 countries who entered the UK after 2004. The marginally higher odds of observing them in elementary occupations relative to the base occupation, when the sample consisted of immigrants with 0 to 5 years of stay in the UK, may have come from comparing older immigrants with new immigrants as that sample would contain immigrants with longer lengths of stay in the before 2004 period in comparison to the after 2004 period.

### **2.7.2.3 Immigration policy and the occupational outcomes of immigrants: refining the post-treatment period**

The change in immigration policy which led to A8 immigrants being free from any visa restriction to enter the UK and have free access to its labour market came into effect in May 2004. As we are unable to observe in which month an individual entered the UK, we exclude all immigrants who entered the UK in 2004 and re-estimate our coefficients of interest from equations (2.7.4) and (2.7.5) in order avoid any potential errors of classification. Estimates of the coefficients of interest are reported in Table 2.9.

**Table 2.8: Odds ratios estimated from a multinomial logit with professional occupations as base occupation; control group is non-EU immigrants.**

|                                           | Immigrants with at most 5 years of stay in UK |                       | Immigrants with 0 to 2 years of stay in UK |                       |
|-------------------------------------------|-----------------------------------------------|-----------------------|--------------------------------------------|-----------------------|
|                                           | (1)                                           | (2)                   | (3)                                        | (4)                   |
| <b>Semi-professional occupations</b>      |                                               |                       |                                            |                       |
| A8*post2004                               | 1.210<br>(0.598)                              | 1.176<br>(0.544)      | 1.149<br>(0.298)                           | 1.190<br>(0.388)      |
| A8 dummy                                  | 3.763***<br>(4.102)                           | 3.624***<br>(4.447)   | 5.091***<br>(3.686)                        | 4.603***<br>(3.839)   |
| Post 2004                                 | 0.881<br>(-0.863)                             | 0.934<br>(-0.474)     | 0.861<br>(-0.856)                          | 0.875<br>(-0.833)     |
| Education                                 |                                               | 0.838***<br>(-7.730)  |                                            | 0.820***<br>(-8.172)  |
| Age                                       |                                               | 1.019<br>(0.730)      |                                            | 1.016<br>(0.317)      |
| Age arrived                               |                                               | 0.970<br>(-0.965)     |                                            | 0.964<br>(-0.711)     |
| Sex                                       |                                               | 0.774<br>(-0.944)     |                                            | 0.716<br>(-1.277)     |
| <b>Skilled manual/routine occupations</b> |                                               |                       |                                            |                       |
| A8*post2004                               | 1.331<br>(1.160)                              | 1.326<br>(1.271)      | 0.870<br>(-0.337)                          | 0.918<br>(-0.214)     |
| A8 dummy                                  | 5.127***<br>(5.730)                           | 4.876***<br>(6.889)   | 10.306***<br>(6.122)                       | 9.072***<br>(6.689)   |
| Post 2004                                 | 1.078<br>(0.852)                              | 1.168*<br>(1.810)     | 1.083<br>(0.945)                           | 1.142<br>(1.431)      |
| Age                                       |                                               | 0.798***<br>(-11.620) |                                            | 0.785***<br>(-10.785) |
| Age arrived                               |                                               | 0.998<br>(-0.042)     |                                            | 1.088<br>(1.179)      |
| Education                                 |                                               | 0.978<br>(-0.520)     |                                            | 0.886<br>(-1.603)     |
| Sex                                       |                                               | 0.402***<br>(-8.651)  |                                            | 0.396***<br>(-8.573)  |
| <b>Elementary occupations</b>             |                                               |                       |                                            |                       |
| A8*post2004                               | 1.579<br>(1.464)                              | 1.441<br>(1.294)      | 1.145<br>(0.280)                           | 1.160<br>(0.310)      |
| A8 dummy                                  | 7.863***<br>(5.155)                           | 7.962***<br>(5.739)   | 14.783***<br>(5.202)                       | 13.110***<br>(5.300)  |
| Post 2004                                 | 1.085<br>(0.932)                              | 1.140<br>(1.508)      | 1.115<br>(1.061)                           | 1.148<br>(1.290)      |
| Education                                 |                                               | 0.760***<br>(-12.794) |                                            | 0.752***<br>(-12.531) |
| Age                                       |                                               | 0.930*<br>(-1.881)    |                                            | 0.979<br>(-0.301)     |
| Age arrived                               |                                               | 1.056<br>(1.358)      |                                            | 0.981<br>(-0.247)     |
| Sex                                       |                                               | 0.700*<br>(-1.705)    |                                            | 0.685<br>(-1.515)     |
| Observations                              | 10,547                                        | 10,547                | 5,932                                      | 5,932                 |
| Log Likelihood                            | -13710.37                                     | -13091.06             | -7590.348                                  | -7229.786             |

Statistical significance: \* p<0.1, \*\*p<0.05, \*\*\*p<0.01. All estimations control for region of residence. t-statistic in brackets (these are negative for values less than 1 – i.e.  $\beta$  less than zero). Standard errors clustered by country of birth.

As with Table 2.8, the set of columns on the left of Table 2.9 contain estimates when the samples consist of immigrants with 0 to 5 years of stay in the UK. When one compares the new estimates to those obtained when the sample included immigrants who entered the UK in 2004, we note that the new estimates are marginally higher in magnitude and there also is an increase in the statistical significance of the coefficients. The change in the odds of being in an elementary occupation relative to the base occupation are now 1.9 (or 1.8)

compared to 1.6 (or 1.4) as reported in Table 2.8. The same patterns re-emerge when looking at skilled manual/routine occupations, or semi-professional occupations.

**Table 2.9: Odds ratios estimated from a multinomial logit with professional occupations as base occupation. Sample excludes 2004 arrivals; control group is non-EU immigrants**

|                                           | Immigrants with at most 5 years of stay in UK |                       | Immigrants with 0 to 2 years of stay in UK |                       |
|-------------------------------------------|-----------------------------------------------|-----------------------|--------------------------------------------|-----------------------|
|                                           | (1)                                           | (2)                   | (3)                                        | (4)                   |
| <b>Semi-professional occupations</b>      |                                               |                       |                                            |                       |
| A8*post2004                               | 1.401<br>(0.978)                              | 1.387<br>(1.016)      | 1.350<br>(0.570)                           | 1.448<br>(0.724)      |
| A8 dummy                                  | 3.783***<br>(4.136)                           | 3.668***<br>(4.538)   | 5.172***<br>(3.715)                        | 4.705***<br>(3.899)   |
| Post 2004                                 | 0.747*<br>(-1.730)                            | 0.781<br>(-1.500)     | 0.733<br>(-1.476)                          | 0.745<br>(-1.512)     |
| Education                                 |                                               | 0.832***<br>(-7.554)  |                                            | 0.815***<br>(-8.144)  |
| Age                                       |                                               | 1.007<br>(0.247)      |                                            | 1.005<br>(0.077)      |
| Age arrived                               |                                               | 0.982<br>(-0.589)     |                                            | 0.977<br>(-0.366)     |
| Sex                                       |                                               | 0.749<br>(-1.046)     |                                            | 0.683<br>(-1.492)     |
| <b>Skilled manual/routine occupations</b> |                                               |                       |                                            |                       |
| A8*post2004                               | 1.507<br>(1.380)                              | 1.523<br>(1.527)      | 0.973<br>(-0.058)                          | 1.054<br>(0.114)      |
| A8 dummy                                  | 5.114***<br>(5.724)                           | 4.847***<br>(6.749)   | 10.313***<br>(6.105)                       | 9.092***<br>(6.639)   |
| Post 2004                                 | 0.966<br>(-0.334)                             | 1.048<br>(0.457)      | 0.954<br>(-0.436)                          | 1.022<br>(0.187)      |
| Education                                 |                                               | 0.798***<br>(-10.572) |                                            | 0.784***<br>(-10.276) |
| Age                                       |                                               | 1.000<br>(0.001)      |                                            | 1.069<br>(0.782)      |
| Age arrived                               |                                               | 0.975<br>(-0.541)     |                                            | 0.902<br>(-1.163)     |
| Sex                                       |                                               | 0.377***<br>(-8.120)  |                                            | 0.367***<br>(-8.306)  |
| <b>Elementary occupations</b>             |                                               |                       |                                            |                       |
| A8*post2004                               | 1.878*<br>(1.822)                             | 1.785*<br>(1.813)     | 1.326<br>(0.535)                           | 1.410<br>(0.660)      |
| A8 dummy                                  | 7.725***<br>(5.067)                           | 7.772***<br>(5.638)   | 14.813***<br>(5.170)                       | 13.185***<br>(5.304)  |
| Post 2004                                 | 1.016<br>(0.146)                              | 1.038<br>(0.340)      | 1.004<br>(0.035)                           | 1.032<br>(0.268)      |
| Education                                 |                                               | 0.761***<br>(-12.221) |                                            | 0.754***<br>(-11.653) |
| Age                                       |                                               | 0.937<br>(-1.511)     |                                            | 0.966<br>(-0.436)     |
| Age arrived                               |                                               | 1.047<br>(1.075)      |                                            | 0.999<br>(-0.017)     |
| Sex                                       |                                               | 0.660**<br>(-2.027)   |                                            | 0.633*<br>(-1.868)    |
| Observations                              | 8,373                                         | 8,373                 | 4,851                                      | 4,851                 |
| Log Likelihood                            | -10873.65                                     | -10364.89             | -6192.394                                  | -5886.635             |

Statistical significance: \* p<0.1, \*\*p<0.05, \*\*\*p<0.01. All estimations control for region of residence. t-statistic in brackets (these are negative for values less than 1 – i.e.  $\beta$  less than zero). Standard errors clustered by country of birth.

The last two columns of Table 2.9 report estimates when the sample is further restricted to new immigrants only – i.e. those with 0 to 2 years of stay in the UK,

with individuals who entered the UK in 2004 dropped from the sample. Here, we also see that the new coefficient estimates are only marginally higher than those obtained when the sample included immigrants who entered the UK in 2004. They however still remain very close to 1 (though statistically insignificant), corroborating earlier findings; evidence that the May 2004 change in immigration policy affected the occupational outcomes of A8 immigrants remains thin.

#### **2.7.2.4 Immigration policy and the occupational outcomes of immigrants: alternative estimation methods**

In this section, we use two alternative estimators to examine the robustness of our findings. We start with an ordered logit estimator, assuming that the classification of occupations into 4 skill level categories by the ONS has an ordinal structure. This assumption seems plausible because looking at Table 2.2 (the table with details of which occupation, at the 2 digit SOC2000 level, belongs to which skill group), we see that occupations in skill level 4 (i.e. professional and managerial occupations) category are more likely to require higher levels of education and also remunerate better in comparison to other occupations. Skill level 3 occupations (i.e. associate professional and skill trade occupations) are more likely to require more years of education (and be more remunerated) than skill level 2 occupations (i.e. semi-skill level occupations). Although it might be unclear whether skill level 2 occupations would require more years of education than skill level 1 occupations (i.e. elementary occupations), they are more likely to have higher remunerations. The ordered logit, unlike the multinomial logit we have used earlier, assumes the relationship between each pair is the same.

The second estimator we use is a Linear Probability Model. Here, we keep professional and managerial occupations as a base category, and alternate the second outcome using one of the other three occupations. In this setting, a coefficient estimate of  $\beta_{\text{DID}}$  which is of magnitude close to 0 would provide supportive evidence to our earlier findings. That is the change in policy did not alter the likelihood of choosing professional occupations versus the alternative occupation being considered amongst A8 immigrants.

#### 2.7.2.4.1 Alternative 1: using an ordered logit estimator

The odds ratios from an ordered logit indicate the odds of being observed in a given occupation versus all occupations ranked higher than that occupation. Table 2.10 reports our coefficient estimates of interest. We see from the first three columns that these coefficients are approximately 0.9 when the sample includes individuals with 0 to 5 years of stay in the UK – though they are statistically insignificant. This suggests almost no change in occupational distribution of A8 immigrants who arrived before and after 2004. More interestingly, looking at the last three columns, where the sample is further restricted to those who have been in the UK for 0 to 2 years at the time of the survey, we see that the coefficients of interest are approximately equal to 1 (though statistically insignificant throughout). Here again, we see no evidence that the occupational outcomes of new A8 immigrants who entered the UK following the change in immigration policy was different to that of new A8 immigrants who entered the UK before the change in immigration policy. This provides further evidence on the robustness of our baseline estimates.

#### 2.7.2.4.2 Alternative 2: Using a linear probability estimator

As noted earlier, we keep the base occupation used in the multinomial estimation, which always takes a value of 1, and alternate the second outcome by choosing one of the other three occupations – we give this outcome a value of 0. Table 2.11 reports our coefficient estimates of  $\beta_{DID}$ . Looking at the first three columns, where estimates when the sample contains immigrants with up to 5 years of stay in the UK are reported, we see that the coefficient on the policy change variable ranges between - 0.027 and - 0.01. This suggests lack of evidence for change in the likelihood of choosing an alternative occupation relative to professional and managerial occupations amongst A8 immigrants as a result of the policy change. Now let's turn to the last three columns of Table 2.11 where we report the coefficient estimates of interest when the sample is further restricted to those who had been in the UK for 0 to 2 years – i.e. only new immigrants. We see that the magnitude of the coefficients ranges between 0.027 and 0.036 (and statistical insignificant here too). It is the coefficient on the A8 dummy that always remains statistically significant and of important economic magnitude. A8 immigrants

appear to be less likely to be in professional occupations relative to other immigrants by over 30 percentage points depending on the occupation.

**Table 2.10: Odd ratios estimated from an ordered logit; control group is non-EU immigrants**

|             | Immigrants with at most 5 years of stay in UK |                      | Immigrants with 0 to 2 years of stay in UK |                      |
|-------------|-----------------------------------------------|----------------------|--------------------------------------------|----------------------|
|             | (1)                                           | (2)                  | (3)                                        | (4)                  |
| A8*post2004 | 0.850<br>(-1.561)                             | 0.874<br>(-1.421)    | 1.018<br>(0.142)                           | 1.003<br>(0.022)     |
| A8 dummy    | 0.356***<br>(-5.847)                          | 0.359***<br>(-6.122) | 0.271***<br>(-6.772)                       | 0.292***<br>(-6.403) |
| Post 2004   | 0.918<br>(-1.349)                             | 0.880**<br>(-2.113)  | 0.899<br>(-1.368)                          | 0.870*<br>(-1.658)   |
| Education   |                                               | 1.166***<br>(9.453)  |                                            | 1.165***<br>(8.148)  |
| Age         |                                               | 1.038<br>(1.466)     |                                            | 1.002<br>(0.065)     |
| Age arrived |                                               | 0.976<br>(-0.950)    |                                            | 1.023<br>(0.573)     |
| Sex         |                                               | 1.451***<br>(3.418)  |                                            | 1.410***<br>(2.713)  |
| N           | 10,547                                        | 10,547               | 5,932                                      | 5,932                |
| LL          | -13850.36                                     | -13397.16            | -7693.618                                  | -7438.9              |

Statistical significance: \* p<0.1, \*\*p<0.05, \*\*\*p<0.1. All estimations control for region of residence. t-statistic in brackets (these are negative for values less than 1 – i.e.  $\beta$  less than zero). Standard errors clustered by country of birth.

**Table 2.11: Average marginal effects from LPM with professional occupations as base occupation; control group is non-EU immigrants**

|             | Immigrants at most 5 years of stay in UK |                          |                       | Immigrants with 0 to 2 years of stay in UK |                          |                       |
|-------------|------------------------------------------|--------------------------|-----------------------|--------------------------------------------|--------------------------|-----------------------|
|             | Semi-prof.                               | Skilled manual / routine | Elementary            | Semi-prof.                                 | Skilled manual / routine | Elementary            |
| A8*post2004 | -0.015<br>(-0.281)                       | -0.027<br>(-0.916)       | -0.010<br>(-0.218)    | 0.007<br>(0.093)                           | 0.036<br>(0.917)         | 0.030<br>(0.588)      |
| A8 dummy    | -0.274***<br>(-4.960)                    | -0.276***<br>(-6.628)    | -0.404***<br>(-6.461) | -0.333***<br>(-4.974)                      | -0.354***<br>(-8.453)    | -0.453***<br>(-7.054) |
| Post 2004   | 0.012<br>(0.411)                         | -0.019<br>(-1.074)       | -0.014<br>(-0.756)    | 0.023<br>(0.727)                           | -0.010<br>(-0.544)       | -0.013<br>(-0.599)    |
| Education   | 0.039***<br>(9.983)                      | 0.042***<br>(13.297)     | 0.047***<br>(12.134)  | 0.043***<br>(11.382)                       | 0.044***<br>(11.786)     | 0.048***<br>(9.776)   |
| Age         | -0.000<br>(-0.013)                       | 0.001<br>(0.139)         | 0.013*<br>(1.857)     | -0.001<br>(-0.080)                         | -0.016<br>(-1.312)       | -0.001<br>(-0.078)    |
| Age arrived | 0.003<br>(0.525)                         | 0.003<br>(0.387)         | -0.010<br>(-1.382)    | 0.005<br>(0.490)                           | 0.022*<br>(1.769)        | 0.007<br>(0.563)      |
| Sex         | 0.089*<br>(1.679)                        | 0.173***<br>(6.638)      | -0.001<br>(-0.041)    | 0.111**<br>(2.207)                         | 0.173***<br>(6.640)      | -0.007<br>(-0.256)    |
| N           | 4,815                                    | 5,410                    | 4,660                 | 2,568                                      | 2,984                    | 2,712                 |
| Adj. R2     | 0.15                                     | 0.26                     | 0.35                  | 0.18                                       | 0.30                     | 0.39                  |

Statistical significance: \* p<0.1, \*\*p<0.05, \*\*\*p<0.1. All estimations control for region of residence. t-statistic in brackets. Standard errors clustered by country of birth

### 2.7.3 Further robustness check: sensitivity to identity of the control group

How sensitive are our estimates to the identity of the control group? To answer the question, we re-estimate equations (2.7.4) and (2.7.5) but using immigrants

from EU14 countries as our new control group. Columns (1) and (2) of Table 2.12 report estimates when the sample consists of immigrants with at most 5 years of stay in the UK. Comparing these estimates to those obtained when the control group was non-EU immigrants, we see that the odds of being in skilled manual/routine occupations relative to professional occupations are around 2, as opposed to around 1.3 when the control group was non-EU immigrants. We observe the same for elementary occupations; the odds of being in an elementary occupation relative to managerial and professional occupation is also around 2 as opposed to around 1.4 when the control group was non-EU immigrants. The odds of being in a semi-professional occupation relative to a professional occupation remain relatively the same compared to those obtained when the control group was non-EU immigrants, at 1.3.

Columns (4) and (5) of Table 2.12 contain estimates when the sample is restricted to new immigrants only – i.e. those who have been in the UK for 0-2 years. Here too, the odds of being in a semi-professional occupation relative to a professional occupation remain relatively the same compared to those obtained when the control group was non-EU immigrants. The odds of being in skilled manual/routine occupations relative to professional occupations are around 1.4, as opposed to around 0.9 when the control group was non-EU immigrants. The odds of being in an elementary occupation relative to a managerial or professional occupation is now around 2.3 as opposed to around 1.1 when the control group was non-EU immigrants. However, none of these new coefficients estimates are statistically significant, even at the 10 per cent level except for the estimate of interest for elementary occupation in column (5).

Columns (3) and (6) report estimates when equation (2.6.4) is extended to include average income levels in the immigrant's country of origin. We now observe a drop in the estimates to more comparable magnitudes as those obtained when the control group was non-EU immigrants. These estimates are now closer to 1 when the sample consists of immigrants with 0 to 2 years in the UK, except the estimate for elementary occupation which reduces to just 1.7 (although now statistically insignificant). Hence, with EU immigrants as an alternative control group largely corroborate earlier findings; what emerges here is the importance of income levels in the immigrant's origin country as a control variable, mainly because unlike non-EU countries which would mostly have comparable income

levels to A8 countries, EU countries may have much higher levels of income in comparison to A8 countries. We shall discuss the role of income levels at the source country on the occupational outcomes of immigrants in the section that follows, but we note that this result further support our preference for immigrants from non-EU countries as the most appropriate control group.

**Table 2.12: Odds ratios estimated from a multinomial logit with professional occupations as base occupation; control group is immigrants from EU14 countries**

|                                           | Immigrants at most 5 years of stay in UK |                       |                     | Immigrants with 0 to 2 years of stay in UK |                       |                    |
|-------------------------------------------|------------------------------------------|-----------------------|---------------------|--------------------------------------------|-----------------------|--------------------|
|                                           | (1)                                      | (2)                   | (3)                 | (4)                                        | (5)                   | (6)                |
| <b>Semi-professional occupations</b>      |                                          |                       |                     |                                            |                       |                    |
| A8*post2004                               | 1.425<br>(1.083)                         | 1.304<br>(0.931)      | 1.335<br>(0.902)    | 1.156<br>(0.287)                           | 1.103<br>(0.205)      | 1.048<br>(0.096)   |
| A8 dummy                                  | 4.849***<br>(5.210)                      | 4.550***<br>(7.066)   | 1.367<br>(0.611)    | 6.874***<br>(4.372)                        | 5.929***<br>(4.623)   | 1.145<br>(0.238)   |
| Post 2004                                 | 0.954<br>(-0.434)                        | 1.052<br>(0.431)      | 1.154<br>(1.607)    | 1.193<br>(0.725)                           | 1.434<br>(1.586)      | 1.803**<br>(2.543) |
| Education                                 |                                          | 0.778***<br>(-7.982)  |                     |                                            | 0.763***<br>(-7.682)  |                    |
| Age                                       |                                          | 0.935**<br>(-2.157)   |                     |                                            | 0.930<br>(-0.849)     |                    |
| Age arrived                               |                                          | 1.043<br>(1.154)      |                     |                                            | 1.042<br>(0.462)      |                    |
| Sex                                       |                                          | 1.419*<br>(1.666)     |                     |                                            | 1.326<br>(1.300)      |                    |
| <b>Skilled manual/routine occupations</b> |                                          |                       |                     |                                            |                       |                    |
| A8*post2004                               | 2.169***<br>(2.747)                      | 2.221***<br>(3.561)   | 1.925**<br>(2.313)  | 1.399<br>(0.706)                           | 1.603<br>(1.083)      | 1.214<br>(0.420)   |
| A8 dummy                                  | 6.666***<br>(6.003)                      | 6.651***<br>(8.844)   | 1.145<br>(0.240)    | 12.374***<br>(6.194)                       | 10.404***<br>(6.382)  | 1.934<br>(1.035)   |
| Post 2004                                 | 0.681***<br>(-2.982)                     | 0.709**<br>(-2.362)   | 0.916<br>(-0.597)   | 0.729<br>(-1.572)                          | 0.822<br>(-1.026)     | 1.159<br>(0.642)   |
| Age                                       |                                          | 0.735***<br>(-13.932) |                     |                                            | 0.728***<br>(-13.798) |                    |
| Age arrived                               |                                          | 0.823***<br>(-7.371)  |                     |                                            | 0.864<br>(-1.165)     |                    |
| Education                                 |                                          | 1.158***<br>(5.818)   |                     |                                            | 1.093<br>(0.710)      |                    |
| Sex                                       |                                          | 0.366***<br>(-7.676)  |                     |                                            | 0.360***<br>(-6.865)  |                    |
| <b>Elementary occupations</b>             |                                          |                       |                     |                                            |                       |                    |
| A8*post2004                               | 2.706***<br>(3.139)                      | 2.498***<br>(3.535)   | 2.233***<br>(2.612) | 2.260<br>(1.593)                           | 2.418*<br>(1.766)     | 1.789<br>(1.222)   |
| A8 dummy                                  | 8.060***<br>(3.558)                      | 9.016***<br>(4.656)   | 0.417<br>(-1.028)   | 12.584***<br>(4.021)                       | 11.129***<br>(4.052)  | 0.631<br>(-0.454)  |
| Post 2004                                 | 0.641***<br>(-5.398)                     | 0.680***<br>(-4.042)  | 1.077<br>(0.613)    | 0.603***<br>(-2.996)                       | 0.704**<br>(-2.207)   | 1.287<br>(0.992)   |
| Education                                 |                                          | 0.687***<br>(-10.156) |                     |                                            | 0.684***<br>(-10.464) |                    |
| Age                                       |                                          | 0.768***<br>(-8.613)  |                     |                                            | 0.812<br>(-1.587)     |                    |
| Age arrived                               |                                          | 1.255***<br>(7.530)   |                     |                                            | 1.161<br>(1.093)      |                    |
| Sex                                       |                                          | 0.462***<br>(-4.898)  |                     |                                            | 0.459***<br>(-3.263)  |                    |
| Observations                              | 4,878                                    | 4,878                 | 4,857               | 3,011                                      | 3,011                 | 2,991              |
| Log Likelihood                            | -5977.00                                 | -5520.302             | -5835.787           | -3592.347                                  | -3331.672             | -3508.281          |

Statistical significance: \* p<0.1, \*\*p<0.05, \*\*\*p<0.01. All estimations control for region of residence. t-statistics in bracket (these are negative for values less than 1 – i.e.  $\beta$  less than zero). Standard errors clustered by country of birth. Columns (3) and (6) contains estimates with income at source country as a control in equation (2.7.4)

## 2.8 Discussion

This section starts by examining two driving factors that are associated with observing A8 immigrants being concentrated in low level occupations: the role of duration of stay in the UK and the role of income levels in the migrant's home country. We then turn to a number of issues relating to our data and the implications for our results. In particular, we discuss the implications of temporary and seasonal migration for our findings, followed by the lack of language proficiency information in our data and hence our analysis.

- *The importance of duration of stay in the UK for occupational outcomes*

As with the case of earnings, one may expect immigrants to move up the occupational ladder as they acquire host country-specific skills (see Chiswick, 1978; and Chiswick et al (2005)). The patterns observed in Figure 2.1 are in line with this. Here, we examine this in a regression framework, controlling for individual characteristics. We focus on A8 immigrants, looking at all those who started their current employment after May 2004. The starting estimated equation is as follows:

$$\log \left[ \frac{P_j}{P_k} \right] = \beta_{j0} + \beta_{j1} \text{duration} + \partial_{j1} \text{age} + \partial_{j2} \text{education} + \partial_{j3} \text{sex} \quad (2.8.1)$$

The second column of results in Table 2.13 presents the odds ratios of being in a given occupation relative to a professional occupation for immigrants with a given length of stay in the UK. Here, the sample comprises A8 immigrants only, and the comparison group is those who have been in the UK for 0 to 2 years. We see that the odds ratio of being in an elementary occupation relative to a professional occupation drops sharply after 3 to 5 years of stay in the UK to only 0.34. The corresponding odds ratios are respectively 0.20 and 0.023 for immigrants who have been in the UK for 6 to 10 years and those with over 10 years in the UK. Similar patterns emerge when we look at skilled manual/routine occupations although at lesser magnitude. We also see that in comparison to immigrants with 0 to 2 years in the UK, immigrants with 3 to 5 or 6 to 10 years in the UK have comparable odds ratios of being in a semi-professional occupation relative to a professional occupation. On the other hand, compared

with immigrants with 0 to 2 years of stay in the UK, the odds ratio of being in a semi-professional occupation relative to a professional occupation was about 0.35 for immigrants with over 10 years of stay in the UK. This suggests an improvement in occupational outcomes with time spent in the UK. In summary, immigrants with longer durations of stay in the UK exhibit better occupational outcomes, even after controlling for individual characteristics.

- *The importance of income levels in the home country of the immigrant*

This section compares the occupational outcomes of A8 immigrants to those of natives, accounting for the role of income levels in the home countries of immigrants. Hence, we extend our sample to include UK-born individuals and have income levels as additional control in equation (2.8.1) in a separate analysis.

The remaining columns of Table 2.13 contain estimates of interest when we compare the occupational distribution of A8 immigrants with different lengths of stay in the UK to the occupational distribution of natives. We assign natives “zero years of stay” in the UK, and as such the estimates in the table report how the odds ratios of being in a given occupation relative to a professional occupation for A8 immigrants compare with those for natives as the stay of immigrants in the UK lengthens. Column (3) contains estimates when we control for average disposable income levels in the individual’s home country, adjusted by PPS (i.e. purchasing power standard)<sup>11</sup> unlike those reported in column (2). We see here too suggestive signs that immigrants move up the occupation ladder with time spent in the country: the gap between the odds ratio of the probability of being in an elementary occupation and the probability of being in a professional occupation for A8 immigrants and natives reduces drastically with time spent in the UK.

Without controlling for income levels in the immigrant’s country of origin, immigrants with 0 to 2 years of stay in the UK are 36 times more likely to be in an elementary occupation than a professional occupation when compared with natives. This drops to 15 for those with 6 to 10 years of stay in the UK and below 2 for those with over 10 years of stay in the UK. When income levels are now included in the estimation, the new estimates are reduced to half the

---

<sup>11</sup>Source: NAMA\_R\_EHHINC2, data downloaded from Eurostat on 21/06/15

magnitude of the earlier estimates. Immigrants with over 10 years of stay in the UK are now as likely as the natives to be in elementary occupations relative to professional occupations. Similar patterns also emerge when we look at skilled manual/routine occupations. Here, however, immigrants with over 10 years of stay in the UK appear slightly less likely to be in these occupations relative to professional occupations in comparison to natives. We also see from column (3) that by 6 to 10 years in the UK, A8 immigrants are less than twice as likely as natives to be in a semi-professional occupation relative a professional occupation. After 10 years in the UK, albeit with no statistical significance, A8 immigrants now appear more likely to be in professional occupations relative to semi-professional occupations when compared with natives. These drastic changes would suggest income levels at the source country of the immigrant play an important role in their occupational outcomes, and would corroborate the possibility of occupational downgrading – the incidence of over-education amongst A8 immigrants in the UK, even relative to immigrants from the EU14 countries, has been documented in Campbell (2013).

Boskin's 1974 study on men and women in the USA confirms three hypotheses: workers choose occupations that 1) maximise the discounted present value of potential earnings, 2) entail the lowest training cost, and 3) offer the lowest discounted present value of expected earning forgone due to unemployment. In the presence of institutional constraints for workers, where the host state channels its policy to attract immigrants with the aim of filling shortages in domestic supply of labour for certain occupations, individual could migrate to the host country principally because the gain in utility from doing so exceeds the current utility they derive from living and working in their home country. This benefits both the immigrant and the host country. In the absence of such institutional constraints, the decision to emigrate would still be motivated by the higher expected utility derived from working in the host country in comparison to working in one's home country but the immigrant now has the flexibility of choosing amongst a diverse range of occupations. However, if other constraints beside immigrations rules lead to individuals choosing occupations that give higher utility compared to home, but that yet are not utility maximising, there might be opportunities for Pareto improvement that could result in welfare improvement for both the host country and the immigrant.

**Table 2.13: Odds ratios estimated from a multinomial logit with professional occupations as base occupation. Roles of duration of stay in the UK and income levels in the migrant's home country**

|                                           | Sample of A8<br>immigrants only | Sample of A8 immigrants and natives |                      |
|-------------------------------------------|---------------------------------|-------------------------------------|----------------------|
|                                           | (1)                             | (2)                                 | (3)                  |
| <b>Semi-professional occupations</b>      |                                 |                                     |                      |
| 0-2 years in the UK                       |                                 | 4.649***<br>(25.751)                | 3.641**<br>(2.487)   |
| 3-5 years in the UK                       | 0.644***<br>(-2.811)            | 2.929***<br>(15.032)                | 2.345*<br>(1.873)    |
| 6-10 years in the UK                      | 0.762***<br>(-2.615)            | 2.136***<br>(3.267)                 | 1.713*<br>(1.817)    |
| 10+ years in the UK                       | 0.353**<br>(-2.284)             | 0.520<br>(-1.345)                   | 0.416<br>(-1.030)    |
| <b>Skilled manual/routine occupations</b> |                                 |                                     |                      |
| 0-2 years in the UK                       |                                 | 8.671***<br>(35.475)                | 3.331<br>(1.519)     |
| 3-5 years in the UK                       | 0.509***<br>(-7.088)            | 5.174***<br>(14.581)                | 2.169<br>(0.998)     |
| 6-10 years in the UK                      | 0.429***<br>(-2.838)            | 3.085***<br>(8.581)                 | 1.298<br>(0.355)     |
| 10+ years in the UK                       | 0.098***<br>(-11.314)           | 0.942<br>(-0.342)                   | 0.400<br>(-1.130)    |
| <b>Elementary occupations</b>             |                                 |                                     |                      |
| 0-2 years in the UK                       |                                 | 36.132***<br>(40.180)               | 18.589***<br>(2.761) |
| 3-5 years in the UK                       | 0.342***<br>(-7.841)            | 15.092***<br>(18.805)               | 8.251**<br>(2.089)   |
| 6-10 years in the UK                      | 0.202***<br>(-5.413)            | 5.813***<br>(5.425)                 | 3.186<br>(1.300)     |
| 10+ years in the UK                       | 0.023***<br>(-4.517)            | 1.740<br>(1.307)                    | 0.960<br>(-0.034)    |
| <b>Additional controls</b>                |                                 |                                     |                      |
| Education                                 | Yes                             | Yes                                 | Yes                  |
| Age                                       | Yes                             | Yes                                 | Yes                  |
| Sex                                       | Yes                             | Yes                                 | Yes                  |
| Income level in country of origin         | No                              | No                                  | Yes                  |
| Observations                              | 4,381                           | 71,593                              | 71,593               |
| LL                                        | -4800.738                       | -90076.73                           | -90074.77            |

Statistical significance: \* p<0.1, \*\*p<0.05, \*\*\*p<0.1. Standard errors clustered by country of birth (these are negative for values less than 1 – i.e.  $\beta$  less than zero). All estimations control for region of residence and year of survey. t-statistics in bracket

- *The implications of temporary and return migration patterns for our findings*

Migration is generally not a one way journey. A large number of immigrants return to their home countries or continue their journey to other destinations. In the context of the UK, Dustmann and Weiss (2007) provide evidence that amongst cohorts of immigrants who entered the country between 1992 and 2002 and observed a year after arrival, 40% of males and 32 % of females left within

five years. In the case of immigrants from A8 countries, the relaxation of immigration policy entailed lower migration cost. Because of the geographic proximity of A8 countries to the UK and the low cost of transportation, the rate of return would likely be much higher amongst A8 immigrants in the after 2004 period in comparison to the before 2004 period. However, at the same time, returnees would possibly return at a later date (Blanchflower and Shadforth, 2009), a migratory pattern more pertinent to seasonal migrants.

The implications of return migration for our results would depend on the nature of selection among returnees. The empirical evidence suggests that as much as returnees can be negatively selected (Wahba, 2015), they can also be positively selected (Rooth and Saarela, 2007). Temporary and seasonal migrants differ in their behaviours in comparison to permanent migrants due to difference in the incentives the two groups face. Dustmann and Weiss (2007) suggest return migration is driven by 1) preference for home country's consumption, 2) higher purchasing power of host country's currency in the home country, and 3) higher return to human capital acquired abroad upon return in the home country. These factors affect the consumption/saving behaviour of temporary and seasonal migrants as well as their labour supply decisions. Some may see migration as an opportunity to accumulate capital to start their own enterprises in the home country (see Dustmann and Kirchamp (2002) and Mesnard (2004)), and non-entrepreneurs would often have higher remuneration upon return home due to the human capital acquired abroad (Wabha, 2015), hence the accumulation of human capital could be a motivation for this group. As points out Clark and Drinkwater (2008), one may observe the following among temporary migrants:

1. They are likely to have lower reservation wages;
2. Saving and reasons of remitting may affect the type of jobs they are willing to take; and
3. Non-economic considerations such as learning English, exposure to different culture, etc. would affect their employment and occupational choices. As such, they are likely to downgrade their occupations because occupational prestige is not important to them.

More generally, our estimates of the odds ratios of observing A8 migrants who entered the UK after 2004 in low skilled occupations would be downward

(upward) biased if migration inflows from A8 countries became characterised by a high proportion of temporary migrants after May 2004 who returned to their home countries were negatively (positively) selected; the reverse holds for estimates of the odds of observing them in high skilled occupations. But, if for reasons unrelated to their skill levels, these temporary migrants “intentionally” undergo (further) occupation downgrading, this may still be attributed to the change in the UK’s immigration policy. The implications are different, however: the change in the UK’s immigration policy would have led to a change in the (behavioural) characteristics, and not the skill composition of these immigrants.

We also note that the survey data we have used in this chapter have some limitation. Seasonal migrants move between their home country and the host country, and household surveys do not generally capture these migrants because they might be in their home countries at the time of the survey. Because the LFS is conducted on a quarterly basis, this provides some appeasement for such concerns, however. Nevertheless, we are mindful that if the proportion of seasonal migrants from A8 countries disproportionately increased after 2004, given that these individuals tend to be employed in skilled manual occupations and elementary occupations, our estimate of impact of the change in policy on the odds ratios of observing A8 migrants who entered the UK after 2004 in these occupations would be downward biased. Akin to this concern is the fact that because (very) short-term migrants tend to live in communal establishments such as guest houses, the LFS would not capture these migrants (Saleheen and Shadforth, 2006). As notes Blanchflower and Shadforth (2009), the proportion of migrants in these establishments is very low (i.e. 2 per cent), therefore the bias resulting from this would likely to be negligible.

- *The role of language proficiency and its implications for our findings*

The selection of immigrants through visa policy takes into account additional observable characteristics to the policy maker such as English proficiency, experience in particular fields or trade, income levels, etc. We are unable to observe these characteristics in the data at hand for most of the sample. We are mindful that if these other characteristics systematically differ amongst immigrants who entered the UK before and after the change in visa policy, and are positively/negatively associated with their occupational outcomes in the UK,

our results will be biased. Nevertheless, one would expect these characteristics to be highly correlated with educational attainment levels and the age of individuals at entry in the UK, which would negate the extent of these biases, or at least reduce them.

We found that the age at arrival in the UK and educational attainment levels amongst A8 immigrants, both observable measures of human capital, did not change much. Additionally, on one hand, one would be inclined to conjecture that the before 2004 arrivals may have had better language skills than the after 2004 arrivals because of the selective immigration policies in place. On the other hand, English has become a common second language on curriculums in many country since the 1990s, which the after 2004 arrivals generation would have benefitted from, unlike their before 2004 arrival counterparts – Okolski and Salt (2014) provides a discussion on the issue for Polish immigrants, the dominant group amongst A8 immigrants. The question of how other characteristics of A8 immigrants that would have been observable to the policy maker before 2004 changed after 2004 would be interesting to explore in further research.

## 2.9 Conclusion

Using a quasi-natural experiment, the chapter has examined the role of immigration policy on the occupational outcomes of immigrants. Looking at a sample of immigrants who entered the UK shortly before and after 2004, the findings suggest that the change in the UK's 2004 immigration policy led to a slight increase, although not statistically robust, in the odds of observing A8 immigrants who entered the UK after 2004 in elementary occupations relative to professional occupations. The drastic shift in occupational distributions of A8 immigrants after the change in immigration policy appears to be partly driven by the fact that most A8 immigrants after 2004 were new in the country. Measures of human capital such as age at arrival in the UK and educational attainment levels amongst A8 immigrants did not change much too. This finding is consistent with the theoretical prediction that reducing migration cost may not necessarily alter the skill composition of migrants, especially if this cost is constant or the extent of its heterogeneity across incoming migrants is relatively small. A clear theoretical prediction of the impact of the annulment of

immigration restriction is an increase in the flow of incoming migrants, which has been observed in the case of A8 immigrants.

The data suggest that A8 immigrants with longer durations of stay in the UK have improved occupational outcomes. This remains the case after controlling for individual characteristics and other covariates in a regression setting. The occupational attainment gap between A8 immigrants and UK-born individuals reduces drastically, however, once one controls for income level in the source country of the immigrant; those with 10 years of stay in the UK now appear to have similar or better occupational outcomes relative to natives.

To the best of our knowledge, this study is the first to examine the effect of a complete relaxation of immigration policy on the occupational outcomes of immigrants. In light of the prospect of new countries joining the EU, this study could be of interest to policy makers. It could indeed also be of interest to those outside the EU who may wish to engage in freedom of labour movement agreements. Additionally, the implementation of restrictive immigration policies are costly to the policy maker, employers as well as migrants. Given the empirical evidence showing high rates of employment amongst A8 immigrants after 2004 and that they are more likely to be net contributors to the public purse (Dustmann and Frattini, 2014), our findings raise the question about how effective and important restrictive immigration policies are for the economy.

It would be interesting to examine how the change in policy affected the occupational outcomes of different cohorts of A8 immigrants who entered the UK after 2004, and their assimilation process too. We leave this on our agenda for future research.

## **2.10 Appendix**

### **2.10.1 Requirements for registering on the Worker Registration Scheme (WRS)**

The requirements for the checks were straightforward, and any genuine A8 applicant ought to have been successful by law:

- (4) Unless the applicant has been issued with a registration card under paragraph (5), the application shall be accompanied by –
  - (a) a registration fee of £50;
  - (b) two passport size photographs of the applicant;
  - (c) the applicant's national identity card or passport issued by the applicant's state;
  - (d) a letter from the employer concerned confirming that the applicant began working for the employer on the date specified in the application.
- (5) In the case of an application by an applicant who has not yet been issued a registration card under this paragraph, the Secretary of State shall, where is satisfied that the application is made in accordance with the regulation and that the applicant –
  - (e) is an accession State worker requiring registration; and
  - (f) began working for the employer on the date specified in the application,send the applicant a registration card and a registration certificate authorising the worker to work for the employer specified in the application, and shall return the applicant's national identity card or passport.
- (6) In the case of any application, the Secretary of State shall, if is satisfied as mentioned in paragraph (5), send the applicant a registration certificate authorising the worker to work for the employer specified in the application.

Taken from "The Accession (Immigration and Worker Registration) Regulation 2004, Part 3, Registration card and certificate"

**Table 2.14: Difference in means of characteristics before and after 2004, by immigrant group. Only immigrants who entered the UK between 2001 and 2006. Sample excludes 2004 arrivals**

|                    | Age at arrival in the UK, migrants from |                       |                       | Age left-continuous full time education, migrants from |                       |                       |
|--------------------|-----------------------------------------|-----------------------|-----------------------|--------------------------------------------------------|-----------------------|-----------------------|
|                    | A8                                      | Non-EU14              | EU14                  | A8                                                     | Non-EU14              | EU14                  |
| After 2004 arrival | 0.726<br>(1.342)                        | 0.616**<br>(2.063)    | -0.342<br>(-0.441)    | -0.220<br>(-1.087)                                     | 0.134<br>(0.922)      | 1.079**<br>(2.843)    |
| Mean               | 27.683***<br>(41.824)                   | 30.581***<br>(82.929) | 29.888***<br>(33.834) | 20.362***<br>(45.418)                                  | 20.595***<br>(94.898) | 20.874***<br>(27.640) |
| N                  | 2,462                                   | 5,921                 | 1,330                 | 2,462                                                  | 5,921                 | 1,330                 |

Statistical significance: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01. t-statistics in bracket. Standard errors clustered by country of birth. Sample comprises those who have lived in the UK for at most 5 years at the time of the survey.

**Table 2.15: Test of differences in changes in characteristics across groups reported in Table 2.14**

|           | Age at arrival in the UK |                  | Age left-continuous full time education |                      |
|-----------|--------------------------|------------------|-----------------------------------------|----------------------|
|           | A8 vs non-EU14           | A8 vs EU14       | A8 vs non-EU14                          | A8 vs EU14           |
| Magnitude | 0.111<br>(0.188)         | 1.069<br>(1.157) | -0.353<br>(-1.480)                      | -1.298**<br>(-3.082) |

Statistical significance: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01. t-statistics in bracket. Standard errors clustered by country of birth. Sample comprises those who have lived in the UK for at most 5 years at the time of the survey.

**Table 2.16: Difference in means of characteristics before and after 2004, by immigrant group. Stock of immigrants in the UK between 2001 and 2006**

|                    | Age at arrival in the UK, migrants from |                       |                       | Age left-continuous full time education, migrants from |                        |                       |
|--------------------|-----------------------------------------|-----------------------|-----------------------|--------------------------------------------------------|------------------------|-----------------------|
|                    | A8                                      | Non-EU14              | EU14                  | A8                                                     | Non-EU14               | EU14                  |
| After 2004 arrival | 0.365<br>(0.387)                        | 0.582**<br>(2.433)    | 0.643<br>(1.453)      | -0.523<br>(-1.389)                                     | -0.098<br>(-0.938)     | -0.216<br>(-1.121)    |
| Mean               | 28.020***<br>(37.148)                   | 30.175***<br>(72.844) | 29.218***<br>(39.055) | 20.444***<br>(63.307)                                  | 20.567***<br>(104.628) | 21.275***<br>(37.743) |
| N                  | 1,431                                   | 6,868                 | 1,747                 | 1,431                                                  | 6,868                  | 1,747                 |

Statistical significance: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01. t-statistics in bracket. Standard errors clustered by country of birth. Sample comprises those who have lived in the UK for at most 5 years at the time of the survey.

**Table 2.17: Test of differences in changes in characteristics across groups reported in Table 2.16**

|           | Age at arrival in the UK |                    | Age left-continuous full time education |                    |
|-----------|--------------------------|--------------------|-----------------------------------------|--------------------|
|           | A8 vs non-EU14           | A8 vs EU14         | A8 vs non-EU14                          | A8 vs EU14         |
| Magnitude | -0.218<br>(-0.237)       | -0.278<br>(-0.278) | -0.424<br>(-1.150)                      | -0.307<br>(-0.754) |

Statistical significance: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01. t-statistics in bracket. Standard errors clustered by country of birth. Sample comprises those who have lived in the UK for at most 5 years at the time of the survey.

**Table 2.18: Placebo effect: looking at immigrants who entered the UK between 1998 and 2003; control group is non-EU immigrants**

|                                                   | <i><b>"fake" change in policy</b></i> |
|---------------------------------------------------|---------------------------------------|
| <b>Associate prof. and skill trade occupation</b> |                                       |
| A8 migrants*( 2001-03 entry )                     | 1.290<br>(1.274)<br>[0.872 1.910]     |
| <b>Semi-skill and trade occupations</b>           |                                       |
| A8 migrants*( 2001-03 entry )                     | 1.084<br>(0.323)<br>[0.664 1.772]     |
| <b>Elementary occupations</b>                     |                                       |
| A8 migrants*( 2001-03 entry )                     | 1.271<br>(1.020)<br>[0.801 2.017]     |
| Year of survey dummies                            | No                                    |
| Observations                                      | 7,695                                 |
| LL                                                | -10247.674                            |

Statistical significance: \* $p < 0.1$ , \*\*\* $p < 0.05$ , \*\*\*\* $p < 0.01$ . Standard errors clustered by country of birth. t-statistics in parentheses and 95% CI in square brackets.

## **Chapter 3: Job-finding networks and migrants' occupational outcomes**

### 3.1 Introduction

“A highly sophisticated network would enable the University of New England with a vacancy for a Mathematical Economist to call the nearest professional office of the employment service, which through its link to a computer in Washington would discover in microseconds a well-trained Mathematical Economist on the Pacific Coast ... Yet even a system with such capabilities would have to struggle for users against the network of personal contacts built up within industries and professions. It is quite possible that the department chairman in New England would prefer, even in making his initial list of prospects, to phone or write to one or two senior Mathematical Economists whose judgement he has learnt to trust.” [Rees, A. (1966)].

A plethora of studies show that the observation made by Rees half a century ago still holds today (see Munshi (2003), Ioannides et al (2004), Giulietti et al (2013), and Cappellari and Tatsiramos (2015)). An explanation for the widespread use of networks of social contacts in the labour market is that “referrals”, which are obtained through one’s network of social contacts, lower search costs for both employers and jobs seekers; Holzer (1988) suggest it is the most efficient and least costly job search method. Relatives, friends and acquaintances would, in most cases, provide job information to the job seeker once they hear of these. More importantly, the job seeker is likely to obtain better information about the nature of the job from the employee making the referral or providing job information, and employers may benefit from the fact that, in the case of referrals, the individual being referred is likely to be a good match (Rees, 1966).

Because networks of social contacts provide job information or employee referrals to the job seeker, jobs found through these networks are likely to be in the same industries or occupations as members of the networks (Bayer et al, 2008; Bentolila et al, 2010; and Patel and Vella, 2013). Identifying who the members of these networks are remains an empirical challenge in studies of social networks and labour market outcomes, however, because the researcher is often not able to observe them directly from the data. Therefore, researchers have tended to use geographical proximity and group affiliation as an alternative measure: social interactions are believed to operate within defined spatial units –

neighbourhoods, metropolitan cities, regions and/or at the firm level – where individuals exchange job information. In the case of immigrants, these networks are commonly defined as co-ethnics in the host country (Munshi, 2003; Andersson et al, 2009; Dustmann et al, 2011; and Patel and Vella, 2013).

This chapter examines the extent to which networks of social contacts of an immigrant job seeker affect the occupational outcome of the job seeker. Following the literature, we define a network based on the country of origin of the immigrant. We are particularly interested in whether the influence of such networks on the occupational outcomes of job seekers varies with the length of stay of immigrants in the host country, as this could entail the presence (or absence) of some form of “social assimilation”. Because the data at hand contain information on the country of birth of respondents, whether employment was found through referral and how long an immigrant has lived in the UK, we are able to empirically examine if co-ethnics still remain the core source of referrals to an immigrant as their duration of stay in the country lengthens.

The most immediate source of information for new immigrants is likely to be fellow co-ethnics in the host country (Munshi, 2003; and Kanas et al, 2012). Hence, while the pool of co-ethnics may serve as a suitable alternative measure of the network of social contacts of new immigrants, it may not necessarily be so for more established immigrants if their networks of social contacts broaden to include non-co-ethnics. This is important because networks of social contacts have been shown to have labour market outcome implications not just for the job seeker but for the general economy too (Patel and Vella, 2013; and Bentolila et al, 2010). Patel and Vella (2013) for instance suggest that the use of networks of social contacts to find employment might explain the growth in the concentration of immigrants in certain occupations across metropolitan cities in the USA. In this example, networks of social contacts are defined as co-ethnics in the host country. A potential long run implication of such a phenomenon is that groups of immigrants would have a monopoly in some occupations, especially if the above relationship holds for older immigrants too. This may arise if the transition between occupations is costly, and/or instead of choosing occupations where they have comparative advantage in skills, the job seeker chooses the prevalent occupation within their network of social contacts because of the

relative lower cost associated with finding employment through the network (Bentolila et al, 2010). Bentolila et al (2010) suggest that such cases could also have adverse implications for aggregate productivity levels in the economy.

Evidence from the UK Labour Force Survey (LFS) suggests that patterns similar to those described in Patel and Vella (2013) may exist in the UK. Amongst immigrants who have been living in the UK for less than 2 years (i.e. new immigrants), the share of new employees who obtained employment through social contacts rose from about 15 per cent to close to 30 per cent between 1998 and 2009. This increase was mirrored, although less steeply, in the share of new immigrants who recently found employment in the most popular occupations amongst non-ethnics already established in their respective occupations. Amongst immigrants who have lived in the UK for over 10 years, on the other hand, the proportion of those who found employment in the most popular occupations amongst fellow co-ethnics was usually below the proportion of recent immigrants who did so. Yet the share of those who found employment through networks of social contacts remained comparable to the corresponding share amongst recent immigrants.

Recent immigrants are arguably less familiar with the functioning of the host country's labour market, and as noted by Kanas et al (2012) and Munshi (2003), therefore their most immediate source of information is fellow co-ethnics in the host country. If immigrants broaden their network of social contacts with time spent in the host country, the occupational distribution of their co-ethnics may have a reduced effect on jobs found through networks of social contacts. As such, the concentration of immigrants from a given country in an occupation might not be sustained as immigrants change and move across jobs during their employment life cycle. This essay will shed light on whether this might indeed be the case.

The existing literature on the UK has focused on the determinants and success of various search methods (Frijters et al, 2005; Battu et al, 2011; and Giulietti et al, 2013). This literature has also examined how the effectiveness of finding employment through networks of social contacts – defined as individuals of the same ethnicity as the job seeker – varies across ethnic groups (Patacchini and Zenou, 2012) and the implications for wages and match quality (Cappellari and

Tatsiramos, 2015). Very little is known about the occupational composition of these networks and their implications for job seekers' occupational outcomes. The present essay fills this gap.

The closest study to ours is Patel and Vella (2013), which look at whether networks of social contacts cause new immigrants to choose the same occupations as their co-ethnics in the USA. Our contribution distinguishes itself from Patel and Vella in the following ways. Firstly, our data contains direct information on whether a job was obtained through a network of social contacts, which allows us to estimate a direct relationship between the likelihood that a job seeker found employment through the network of social contacts in a given occupation and the size of their network of social contacts in that occupation. Secondly, unlike in Patel and Vella (2013) where they focus only on new immigrants, we examine whether networks of social contacts lead both new immigrants and old immigrants to choosing the same occupations as their co-ethnics. If immigrants broaden their network of social contacts with time spent in the country, the occupational distribution of co-ethnics may have lesser implications for jobs found through these networks, which would also suggest some form of "social assimilation". Thirdly, while Patel and Vella (2013) measure the strength of the network in channelling job seekers toward the same occupations as its members, it is not informative as to the types of occupations where networks may play a stronger role. This is because in their analysis only two outcomes are possible: the job seeker either finds employment in the most popular occupation or any other occupation. We estimate a more flexible model with a broader set of possible outcomes, based on nine 1-digit occupational codes. This novel approach enables us to estimate the relationship between the likelihood of an immigrant being observed in each of these nine occupations and the size of their network of social contacts in the occupation. Because a job seeker can receive job information in any occupation where co-ethnics are employed, their network of social contacts may vary across occupations, but would not vary across immigrants from the same home country.

Our findings suggest that co-ethnics already established in their respective occupations have a stronger influence on the occupational outcomes of recent immigrants than those of immigrants with longer durations of stay in the UK.

This may suggest that co-ethnics form the initial social network of contacts of immigrants in the host country, but that they potentially become “socially assimilated” with time spent in the UK by broadening their circles of social contacts.

In what follows, we present a review of the literature in section 2 and some empirical observations from our data in section 3. Section 4 outlines underlining mechanisms that could explain what we observe from the data and lays out some hypotheses. Section 5 outlines our methodology and the hypotheses in section 4 are tested in Section 6. Section 7 concludes the chapter.

### **3.2 Literature Review**

A simple reason why referrals and transmission of job vacancy information are predominant in the labour market is because they reduce asymmetry of information between job seekers and employers: “Employee referrals – the most important informal channel – usually provide good screening for employers who are satisfied with their present workforce. Present employees tend to refer people like themselves, and they may feel that their own reputation is affected by the quality of the referral. ... The informal sources also have benefits to the applicant. He can obtain much more information from a friend who does the kind of work in which he is interested than from an ad in the paper or a counsellor at an employment agency, and he puts more trust in it.” (Rees, 1966). The mechanisms behind this phenomenon and its labour market implications have been formalised in later studies. Generally, individuals are embedded in social groups consisting of employed individuals and job seekers. Group members in employment have access to job information in their respective segments of the labour market, which they transmit to group members who are seeking employment (see Montgomery (1991) and Calvo and Jackson (2004) for example).

### 3.2.1 Defining and measuring networks

Defining group memberships or networks is a challenge in empirical studies due to data limitations. Typically, data does not identify networks (e.g. acquaintances, friends and relations) directly<sup>12</sup>. Researchers have therefore used geographical proximity and group affiliation as alternative measures of these social interactions in order to examine the presence of network effects in the labour market (these include Munshi (2003), Wahba and Zenou (2005), and Bayer et al (2008)). In this stream of the literature, social interactions are believed to operate within defined spatial units such as neighbourhoods, metropolitan cities, regions, etc, and within defined social units such as nationality/country of origin. The researcher clusters individuals by group based on some observed characteristics defining group affiliations and/or geographic proximity. In the case of immigrants, the country of origin of the immigrant has often been used as an additional dimension for the clustering (Munshi, 2003; Andersson et al, 2009; Beaman, 2012; Dustmann et al, 2011; and Patel and Vella, 2013). Networks are generally defined based on the share or number of co-ethnics within these clusters. The relationship between various labour market outcomes – for example employment probability, the likelihood of obtaining employment through referrals, etc. – and the size and/or the “quality” of the network is then examined. A general form of the estimated equation is:

$$y_{ik} = \beta_0 + \beta_1 Network_{ik} + \varepsilon_{ik} \quad (3.2.1)$$

where  $y_{ik}$  denotes the outcome of interest from individual  $i$  in cluster  $k$ .  $Network_{ik}$  denotes the network of  $i$ , and usually measures of the stock or the proportion of network members in the cluster.  $k$  indicates how individuals are clustered, which can be by place of residence, place of work and/or country of birth.

Typically, neither the way employment was obtained nor a measure of kinship is observed from the data, and alternative measures are used for both. One such example is Bayer et al (2008). They evaluate the likelihood that two individuals

---

<sup>12</sup> Exceptions are Cappellari and Tatsiramos (2010, 2015) who use the British Household Panel Survey which contains information on individuals' 3 best friends and their characteristics. Because there is no information on how employment was obtained, their interest lies in the relationship between transition from unemployment to employment and the employment status of one's best friends.

work in the same location given that they reside in the same census block compared with living in nearby blocks. The idea here is that the propensity of two individuals to work in the same location is an indication that one of these individuals provided referral/job information to the other. They find that living in the same block versus living in nearby blocks increases individuals' propensity to work in the same location by 33 per cent. Similarly, in Dustmann et al (2011) there is no direct information on how employment was obtained, nor is there a measure of kinship. Instead, they examine the relationship between the share of the new hires in a firm who are from a given country of origin and the share of pre-existing employees from the same country amongst all pre-existing employees in the firm. Kinship/friendship is believed to be stronger amongst individuals from the same country, and pre-existing employees within the network would likely provide referral to a job seeker from the same place of origin as them when such an opportunity arises in the firm. They find that a 10 percentage point increase in the existing share of employees from an ethnic minority group in a firm leads to a 5.7 percentage point increase in the share of those from the same place of birth amongst all newly hired employees in the firm, an effect attributed to presence of referrals.

A similar paper to Dustmann et al (2011) is Patel and Vella (2013) where the data neither identifies how employment was obtained nor how individuals are interconnected. Here the size of the network of a new immigrant is measured by the within group share of established co-ethnics in a given occupation (with metropolitan area as the unit of analysis), and the interest of the authors lies in examining the likelihood that a new immigrant chooses the most popular occupation amongst established co-ethnics and the employment share of established co-ethnics amongst all co-ethnics in that occupation. They estimate that a 1 percentage point increase in the share of the popular occupation amongst established co-ethnics increases the likelihood that a new immigrant chooses this most popular occupation by 1.2 to 5 percentage points in a given USA metropolitan area.

Studies using co-ethnics as a measure of the network of an immigrant include Munshi (2003), Andersson et al (2009) and Beaman (2012). In Munshi (2003), the size of the network is measured by the proportion of established co-ethnics

(those with 3 or more years of stay in the USA) from the new immigrant's community of origin, who are at the immigrant's locality of residence in the USA. Both Andersson et al (2009) and Beaman (2012) on the other hand use the absolute size of co-ethnics as a measure of network size. In these cases, the size of a network is measured by the stock of co-ethnics in the same locality as the new immigrants, who are already established in the host country. While Munshi (2003) and Andersson et al (2009) look at the relationship between the size of the network and the employment probability of a new immigrant, Beaman (2012) examines how the sizes of different cohorts within the network affect the employment probability of a new cohort of immigrants. In these studies too, the data does not identify whether employment was obtained through referral or not.

Wahba and Zenou (2005) and Patacchini and Zenou (2012) both observe whether employment was obtained through referral, and are therefore able to examine how the role of networks (via helping find employment through referral) may vary across different subgroups of the population. Wahba and Zenou (2005) first illustrate, in a theoretical framework, how reliance on a network to find employment differs across the least educated job-seekers and the well-educated. The size of the network through which referral is acquired or job information is transmitted is measured by the population density of the location where the job seeker resides, and its quality is measured by the employment rate within the network – intuitively, people living in the same locality would interact with one other and exchange job information in the process. The findings suggest that a 10 per cent increase in area density leads to a 2.2 per cent increase in the probability of having found employment through referral, with the effect being more pronounced amongst the illiterates/uneducated and those in blue collar jobs.

Patacchini and Zenou (2012), on the other hand, examine the use of referral across different ethnic groups in the UK. In this study, the job seeker's network of social contacts is defined as individuals of working age amongst their ethnic group who live in a given proximity to them. The findings show that for some ethnic groups, the likelihood of having found employment through referral increases in the employment rate within the network, but at a decreasing rate with distance (measured in travel time) from the network. Amongst other ethnic groups, the employment rate of individuals from the same ethnicity has little

effect on the likelihood of finding employment through referral. They suggest that the network variable used in the paper may not capture the social interactions that lead to referrals for these ethnic groups.

### 3.2.2 Identifying the network effect

A key identification challenge is separating the true network effect from the effect of unobservable characteristics which are shared by network members. This arises because individuals embedded within a social network tend to behave similarly as 1) their behaviour is affected by the same exogenous factor(s); 2) they have similar characteristics; and 3) their individual behaviour depends on the group behaviour. These are respectively known as the exogenous effect, the correlated effect and the endogenous effect. The last effect is the subject of interest. The estimation challenge resulting from these concerns is commonly known as the “reflection” problem, as coined by Manski (1993).

The researcher can often use information in the data to address the reflection problem. A common approach has been to clearly delineate members within the group who provide job information/referral (i.e. the network of social contacts) from members within the group who receive this information (i.e. job seekers) using available information in the data to address concerns about the exogenous effect. A rich set of fixed effects, which often consist of place of work/residence and country of origin fixed effects, is used to address concerns about the correlated effect. The error term in equation (3.2.1) is thereby segmented to account for these fixed effects, and macro level variables are often included in the estimation to account for aggregate factors likely to affect the outcomes of job seekers and those of their respective networks of social contacts.

Dustmann et al (2011) for instance use tenure within a firm as a mean of delineation, and newly employed individuals are assumed to have received referral from pre-existing fellow co-ethnics within the firm. Beaman (2012), on the other hand, group immigrants by their cohort of arrival and examine how the sizes of various older cohorts affect the employment probability of a new cohort. Munshi (2003) classify immigrants with less than 3 years of stay in the USA as new immigrants; established immigrants, i.e. the network, are those who have

been in the USA for more than 3 years. He then examines whether the network has any effect on the employment probabilities of new immigrants. Patel and Vella (2013) use a similar delineation as Munshi (2003). They separate old and new immigrants using information on duration of stay in the USA: new immigrants are defined as those who have lived in the USA for no more than 5 years and established immigrants are defined as those who have lived in the USA for at least 5 years. Established immigrants form the network that provides job information/referral to the new immigrants.

These delineations may, however, not always be sufficient to address the reflection problem and uncover the network effect of interest. In Munshi (2003) for example, where the interest lies in examining whether the network has any effect on the employment probability of new immigrants, it is argued that some macro shocks to the economy may be persistent over time. Past shocks would therefore have affected the current size of the network (i.e. individuals who arrived in the USA prior to the new immigrant's arrival) and would also affect the current employment probability of new immigrants through persistence. This is justified by the responsiveness of the flow (and stock) of Mexican immigrants in the USA to its economic climate. Munshi therefore instruments the size of the network by the lag measure of rainfall in the new immigrant's community of origin. The intuition here is that severe rainfall in the past would have led to higher rate of emigration to USA, increasing the size of co-ethnics in the USA but with no direct implication for the employment probability of a new immigrant. Another example of a study which uses an instrumental variable is Patel and Vella (2013). Despite delineating the pool of network of social contacts and the pool of job seekers using information on duration of stay of the immigrant in the USA, the occupational outcomes of new and old immigrants might be simultaneously determined by the same macro factors because it is not possible to observe when individuals started their current employment from the data used. They therefore instrument the size of a migrant's network by the lag value of established co-ethnics to uncover the network effect of interest.

The present essay is similar to Patel and Vella (2013) as it seeks to examine how the occupational composition of the network of social contacts of an immigrant affects their occupational outcome in the context of the UK. Unlike in Patel and

Vella (2013), however, the data at hand contains information on whether employment was found through referral and how long an immigrant has lived in the UK. As such, we are able to first empirically examine if co-ethnics still remain the core source of referrals to an immigrant as their duration of stay in the country lengthens; it is presently unclear from the existing literature whether reliance on fellow co-ethnics to find employment is a short-term phenomenon amongst new immigrants or persists with time spent in the host country.

In replicating Patel and Vella (2013), we go further to examine if the tendency of networks to cause immigrants to choose the same occupations as co-ethnics changes with length of stay of the immigrant in the host country. One additional contribution of this chapter is that we later depart from examining whether or not a newly found job was in the most popular occupation of their co-ethnics, the method used in Patel and Vella, and use a more flexible estimator which allows for choice and individual attributes. This makes it possible to estimate the relationship between the size of the network in a given occupation and the likelihood that the job seeker found employment in that occupation.

### **3.3 The data**

This study uses data from the LFS, which contains an array of information about labour market outcomes as well as socio-demographic characteristics. The survey is a rotating panel whereby individuals are interviewed for 5 consecutive quarters, with outgoing households replaced with new ones, as mentioned in the precedent chapter. With a sample of about 53,000 to 60,000 private households per quarter, it has been widely used in studies examining the labour market outcomes of immigrants, arguably because it is the survey with the largest sample of immigrants in the UK. The LFS contains information on the place of birth, place of residence, place of work, job search methods, the employment status and occupation of the respondent. As such, it has very often been used to examine the use of networks in the labour market amongst both natives and immigrants (examples include Frijters et al (2005), Sumption (2009), Battu et al (2011), Pattachini and Zenou (2012), and Giulietti et al (2013)).

Given the same individual could be interviewed up to 5 consecutive quarters, we only keep observations when respondents were observed in the survey for the first time. Because information on place of work is only consistently available in every quarter from 1998 onward in our data and respondents cannot be uniquely identified across quarters from 2010 onward, our sample covers the period of 1998 to 2009.

The sample consists of foreign-born individuals between the ages of 21 and 64, excluding those still in full-time education. The LFS provides information on the year in which a foreign-born individual first entered the UK. We use this information in conjunction with the age of the respondent to compute their age at arrival in the UK, and subsequently restrict our sample to individuals who entered the UK at the age of 21 or above. These individuals are likely to be economic migrants, and hence those of interest to the present essay.

An additional attractiveness of the LFS is that it contains information on tenure with one's current employer. We refer to those who have been with their current employer for less than a year as "new employees", and those who have been with their current employer for over a year as "established employees". We first group individuals based on country of birth, and information on region of work is subsequently used to construct the size of one's network of social contact. The network of social contacts of a new employee is measured by the size of fellow co-ethnics who are established employees in their current jobs and work in the same region as the new employee – we take an average value over 5 years to circumvent data limitations. Established employees in their respective occupations are assumed to be able to provide referrals or viable job information to job seekers they know.

In every survey, individuals are asked how they obtained their current employment. The question asked and the responses to choose from are as follows: *Did you get the work that you were doing/were away from the week ending Sunday [date] through...*

1. *replying to job advertisement;*
2. *a jobcentre/job-market or training and employment agency office;*
3. *a careers office;*
4. *a job club;*
5. *a private employment agency or business;*

6. *hearing from someone who worked there;*
7. *a direct application; or*
8. *some other way.*

We consider choosing the 6<sup>th</sup> option as an indication of having found employment through referral (i.e. network of social contacts). Table 3.1 presents the proportion of new employees who found their jobs through their network of social contacts – for immigrants, these proportions are presented by duration of stay in the UK. The sample consists of all those who responded to the question on how they obtained their current employment. Before 2005, this question was asked of those who had found employment within the last 3 months prior to the time of the survey; from 2005 onward, all employees who had found employment within the last 12 months prior to the time of the survey were asked the question. Similarly to Patacchini and Zenou (2012) and Giuliatti et al (2013), we pool data from the before and after 2005 periods<sup>13</sup>.

Firstly, we see that about a quarter of all jobs are found through networks of social contacts. Secondly, despite marginal differences, the overall shares of new jobs found through networks of social contacts remain similar across natives and immigrants as well as across immigrants with different durations of stay in the UK (except those with over 10 years of stay in the UK), suggesting that the use of networks of social contacts to find employment is widespread across all groups in the population.

---

<sup>13</sup> From 2005 onward, the shares of those who obtained employment through referral amongst those who have been with their current employer for 3 months or less are comparable to the shares of those who obtained employment through referral amongst those who have been with their current employer for 3 to 12 months. Additionally, less than 5 percent of those who have been with their current employer for less than 3 months were unemployed 3 months later – see Table 3.11 and Table 3.12 in the appendix.

**Table 3.1: How job was obtained (in %) – new employees, LFS 1998-2009**

|                                 | Through network of social contacts | Through other means |
|---------------------------------|------------------------------------|---------------------|
| All                             | 24.54                              | 75.46               |
| Natives                         | 24.51                              | 75.49               |
| <i>Immigrants with stay of:</i> |                                    |                     |
| Less than 2 years in the UK     | 25.41                              | 74.59               |
| 2 - 5 years in the UK           | 24.24                              | 75.76               |
| 6 - 10 years in the UK          | 23.81                              | 76.19               |
| Over 10 years in the UK         | 28.02                              | 71.98               |

*The sample consists of those who started their current job in the last 3 months prior to the survey before 2005 and those who started their current job in the last 12 months prior to the survey after 2005.*

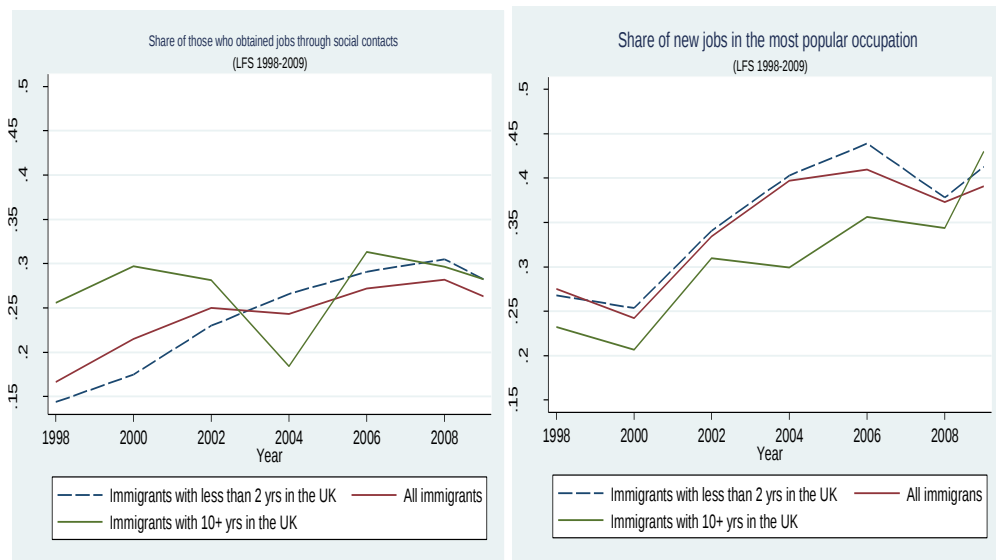
Figure 3.1 reports the share of new employees who found employment through networks of social contacts and the share of new employees in the most popular occupations amongst fellow co-ethnics who are established employees. Note that Figure 3.1 only refers to immigrants. In each graph, the maroon lines represent all immigrants, the green lines represent immigrants who have been in the UK for over 10 years and the blue dashed lines represent immigrants who have been in the UK for less than 2 years. Looking at the graph on the left half of Figure 3.1, we see that the share of all those who found employment through their network of social contacts has gradually been increasing, in particular amongst immigrants who have been in the UK for less than 2 years. Amongst immigrants who have been in the UK for over 10 years, the trend often fluctuates and is somewhat unclear.

A similar trend is observed when looking at the proportion of new employees who found employment in the most popular occupation amongst their co-ethnics who are established employees. Immigrants who have been in the UK for less than 2 years tend to find employment in the most popular occupations of co-ethnics at more than the average rate across all immigrants. On the other hand, the share of new jobs in the most popular occupations amongst fellow co-ethnics has been consistently below the average for immigrants with longer stays in the UK, suggesting immigrants may be less likely to choose the most popular occupation amongst co-ethnics as their duration of stay in the host country lengthens.

Table 3.2 presents the shares of those who found employment through their networks of social contacts by occupation. We observe that in all 9 occupations listed in Table 3.2, a considerable share of individuals found employment

through their networks of social contacts. We also see that the proportion of those who did so varies a lot across occupations; while over a third of those employed in elementary occupations found employment through their network of social contacts, the corresponding share was only about 13 per cent in professional occupations. This suggests that the ability of one's network of social contacts to help one find employment may also vary across occupations. Because the estimator used in Patel and Vella (2013) only allows for two outcomes – i.e. the job seeker either finds employment in the most popular occupation or any other occupation – such heterogeneity would be lost.

**Figure 3.1: share of new jobs found through networks of social contacts and share of new jobs in most popular occupations**



*The sample consists of those who started their current job in the last 3 months prior to the survey before 2005 and those who started their current job in the last 12 months prior to the survey after 2005.*

**Table 3.2: How job was obtained, by occupation (in %) – immigrants only, LFS 1998-2009**

|                                              | <i>Through network of social contacts</i> | <i>Through other means</i> |
|----------------------------------------------|-------------------------------------------|----------------------------|
| Managerial occupations                       | 17.55                                     | 82.45                      |
| Professional occupations                     | 12.11                                     | 87.89                      |
| Associate professional occupations           | 16.51                                     | 83.49                      |
| Clerical and secretarial occupations         | 15.77                                     | 84.23                      |
| Skill trade / craft occupations              | 37.10                                     | 62.90                      |
| Personal service occupations                 | 23.33                                     | 76.67                      |
| Sales and customer service occupations       | 29.00                                     | 71.00                      |
| Process, plants and machine ops. Occupations | 27.16                                     | 72.84                      |
| Elementary occupations                       | 36.11                                     | 63.89                      |

*The sample consists of those who started their current job in the last 3 months prior to the survey before 2005 and those who started their current job in the last 12 months prior to the survey after 2005.*

At this point, we recap the key variables and definitions used:

- *Jobs found through the network of social contacts*: jobs which are obtained by “hearing from someone who worked there”.
- *New employees*: individuals who have been with their current employer for at most one year.
- *Established employees / the network of social contacts*: co-ethnics who were in employment when the job seeker was looking for employment. These are individuals who were in a position to provide job vacancy information or referrals to those who recently found employment. They are those who have been with their current employer for at least a year and work in the same locality (region of work is our main spatial unit) as the new employee. This is measured as an average over 5 years to circumvent data limitations.

### 3.4 Job networks and occupational outcomes

As discussed in the literature review, a key reason for the widespread use of network of social contacts in the labour market is that it reduces search cost for both job seekers and employers (Rees, 1966). A number of authors have formalised Rees’s observations in various theoretical frameworks (Calvo-Armengol and Zenou, 2005; Wahha and Zenou, 2005, Patacchini and Zenou, 2012; Fontaine (2007, 2008); Cahuc and Fontaine, 2009; and Bramoullé and Saint-Paul, 2009). In these frameworks, job search models are extended to incorporate the role social networks play in reducing search cost through transmission of job information and in matching job seekers and employers: job

seekers and employed individuals are embedded in social groups. Job information arrive in the network through those in employment and is passed onto unemployed group members who are seeking employment. It follows from this that the higher the number of employed individuals in the job seeker's social network, the higher their job finding rate would be (Cappellari and Tatsiramos, 2015). Further extensions account for on the job search, whereby the job seeker can either be employed or unemployed (recent examples include Roman Chuhay (2013), Galenianos (2014) and Stupnytska and Zaharieva (2015)).

In this section, we present two hypotheses, which inform our empirical analysis and are motivated by the theoretical framework mentioned above. In deriving these hypotheses our focus is on occupational outcomes, rather than job finding rates and match quality. Studies incorporating occupational outcomes in the general framework outlined above include Horvath (2014), Patel and Vella (2007, 2013), Bentolila et al (2010) and Zaharieva (2016). Bentolila (2010), for example, show social networks may generate occupational mismatch because given they reduce search cost, they may gravitate job seekers toward occupations in which they have less productive endowments. The findings in Horvath (2014), however, suggest that social network can reduce mismatch for particular levels of homophily. Zaharieva (2016), on the other hand, point out that it may be important that one's network of social contacts is diversified to include contacts in occupations in which the job seeker could be less productive. Here, business cycle fluctuations may heterogeneously affect demand for labour in different occupations, hence a diversified network across occupations reduces volatility in income that arises from these demand shocks. Patel and Vella (2007) derive two theoretical predications, which are then supported with empirical evidence using USA survey data: 1) the probability that an individual is observed in an occupation is a function of the size of their network in that occupation; and 2) an individual's reservation wage, hence their observed accepted wage, depends on the size of the individual's network in that occupation. Our hypotheses closely relate to 1). We outlined below a (parsimonious) theoretical framework that leads to these.

Consider an economy with 2 subgroups of workers ( $g = 1, 2$ ) and a labour market with 2 occupations ( $o=1, 2$ ). The two subgroups represent two different places of birth, e.g. natives and migrants. Employed members of each subgroup

can hear about job vacancies in their respective occupations and can relay information about these job vacancies to job seekers in their own subgroup.

A job seeker can find employment either through their network of social contacts (by referral, for example) or through other means. Let  $\lambda(ntk)$  denote the job arrival rate through the network and let  $\lambda(oth)$  denote the job arrival rate through other search methods. Further assume that the job arrival rate through other search methods in a given occupation of the labour market is identical for individuals across subgroups, i.e.  $\lambda(oth) = \lambda$ , but the job arrival rate through the network is a function of the job seeker's network and other attributes specific to their subgroup:  $\lambda_g(ntk) = f(ntk_g, attribute_g, \lambda)$ ; where  $g$  denotes the subgroup the job seeker is a member of, and  $attribute_g$  denotes non-occupational attributes particular to that subgroup.  $ntk_g$  denotes the network that provides referrals or job information to the job seeker. The ratio of the job arrival rate through network and the job arrival rate through other search methods can be expressed as<sup>14</sup>:

$$\frac{\lambda_g(ntk)}{\lambda_g(oth)} = f(ntk_g, attribute_g) \quad (3.4.1)$$

Assuming that the probability of finding a job through the network relative to finding a job through other search methods may be expressed as a function of (3.4.1), one has the following:

$$\frac{\text{Pr(job found through the network)}}{\text{Pr(job found through other methods)}} = f\left(\frac{\lambda_g(ntk)}{\lambda_g(oth)}\right) = f(ntk_g, attribute_g) \quad (3.4.2)$$

In other words:

$$\text{Pr (job obtained through referral)} = f(ntk_g, attribute_g) \quad (3.4.3)$$

For a given occupation  $O$ , it follows that:

$$\text{Pr}_o(\text{job obtained through referral}) = f(ntk_{g_o}, attribute_g) \quad (3.4.4)$$

These equations give rise to our hypotheses:

**Hypothesis 1:** Because employed members of the network relay job vacancy information to job seekers in the network, as they hear of these in their

---

<sup>14</sup> Note the underlying relationship between the probability of finding a job by referrals relative to finding a job through other search methods does not change if we assume that  $\lambda(oth) \neq c$ . What is required is that  $\frac{\lambda_g(ref)}{\lambda_g(oth)} = f(ntk_g, attribute_g)$

respective occupations, the larger the number of employed members of the network (i.e. pool of the source of job information), the higher the job arrival rate through the network would be within the subgroup, and hence the higher the probability of finding employment through the network relative to other methods of search in that segment of the labour market.

**Hypothesis 2:** Referrals and job information from the network are likely to lead to employment in the most popular occupation. This follows from *Hypothesis 1*.

Suppose an individual faces  $O$  alternative choices of occupations, indexed by  $o = 1, 2, \dots, O$ . The utility derived from a given choice may be written as:

$$U_o = D_o + \varepsilon_o \quad (3.4.5)$$

where  $D_o$  is a non-stochastic component of  $U_o$  and  $\varepsilon_o$ , the random component of  $U_o$ , reflects the idiosyncrasies of the individual's taste.

One can rewrite  $D_o$  such that  $D_o = d_o + s_o$ , where  $s_o$  denotes the deterministic component which is a function of job arrival rate through the network of social contacts relative to other means.  $d_o$  denotes the deterministic component that is a function of other characteristics observable to the researcher. Assuming that  $\varepsilon_o$  takes the form of an independent identically distributed Extreme Value Type-1 Distribution, as shown by McFadden (1974), the probability of choosing a given occupation takes the form of:

$$P_o = \frac{e^{d_o + s_o}}{\sum_{o \in O} e^{d_o + s_o}}$$

$$\text{Such that } \frac{P_{o_1}}{P_{o_2}} = \frac{e^{d_{o_1} + s_{o_1}}}{e^{d_{o_2} + s_{o_2}}} \text{ and } \log[P_{o_1}/P_{o_2}] = (d_{o_1} - d_{o_2}) + (s_{o_1} - s_{o_2}) \quad (3.4.6)$$

Assuming  $s_o$  is a positive function of the size of the network in a given occupation as we did earlier,  $\lambda_{o_1}(ntk) > \lambda_{o_2}(ntk)$  implies  $s_{o_1} > s_{o_2}$ . It follows that  $P_{o_1} > P_{o_2}$  holding all else constant. As networks of social contacts reduce job search cost, the higher  $s_o$  is, the lower search cost would be, such that the individual's utility increases in  $s_o$ .

We examine *Hypothesis (1)* and *Hypothesis (2)* for immigrants with different lengths of stay in the UK.

### 3.5 Methodology

In an ideal setting, one would observe job seekers who are randomly allocated into networks of social contacts from whom job information or referrals are obtained. The researcher would also observe the occupations of members of the network as well as the occupations where they provide job information or referrals. Suppose the occupational composition of the different networks exogenously differs across groups; for example, while one group may have a higher proportion of network members in one occupation, the other may have a higher proportion of network members in another occupation. The data would also allow the researcher to observe the country of birth of the job seeker and that of each member of their respective networks, the nature of job information/referrals from these members of the network and the occupation that the job seeker finally chooses. Such an ideal setting is rarely observed in real life.

Our data has the advantage of identifying whether employment was obtained through referral (i.e. from one's network of social contacts), but we do not observe the characteristics of the person(s) who made the referral or the nature of social interaction that led to referral. This problem of missing data is an empirical challenge often faced in the literature. As noted in the literature review, researchers have therefore used geographical proximity and group affiliation as an alternative measure of these social interactions. Individuals are clustered by group based on some observed characteristics defining group affiliations and/or geographical proximity. When the focus of the study is on immigrants, it has generally been assumed that the network of social contacts through which employment was obtained is co-ethnics of the migrant who live or work in close proximity (Dustmann et al, 2011; and Patel and Vella, 2013). We adopt a similar definition of the network because we do not observe who provided job information/referral to the job seeker.

An additional challenge is the “reflection problem”, as mentioned earlier. Manski (1993) uses the analogy of observing the movements of a person in front of a mirror; it soon becomes unclear to an observer whether the individual is causing the observed movements from the mirror or vice versa, as Manski points out. As outlined in the literature review, a clear delineation of the network which

provides job information and who the job seekers are allows one to address key causes of the reflection problem in this context. Because we are able to observe how long the respondents have been with their current employers from the data, the network of social contacts through which employment was obtained is defined as co-ethnics who were in employment when the newly employed migrant was looking for job. These are individuals who would have already been with their current employer, and hence have been able to provide job information or referrals to the newly employed during their job search. This allows us to overcome the exogenous effect. We also use a set of fixed effects consisting of region of residence dummies, country of birth dummies and occupation dummies to control for similar characteristics between the job seeker and network members. This addresses the correlated effect mentioned earlier in section 3.2.

### **3.5.1 The relationship between networks of social contacts and occupational outcomes**

Our first hypothesis is that the likelihood of finding employment through referral in a given occupation increases in the size of the network of social contacts in that occupation. The second hypothesis is that referral and job information from networks of social contacts may lead to employment in the most popular occupation within the network of social contacts. Here, we follow Patel and Vella (2013) who study a similar question in the context of the USA. Only two outcomes are however allowed therein: the dependent variable takes value 1 if employment was found in the most popular occupation, else it takes value 0 – the estimation used is a Linear Probability Model (LPM). While this approach does allow us to examine whether the migrant's network of social contacts, via its occupational composition, affects one's occupational outcome, it does not allow us to gauge the relationship between the size of a migrant's network of contacts in a given occupation and their likelihood of being observed in that occupation. However, this is arguably important because there could be differences in the importance of networks of social contacts in each occupation the job seeker may choose amongst all available alternatives. Because a job seeker can receive job information in any occupation where established co-ethnics are employed, their

network of social contacts may vary across occupations but would not vary across immigrants from the same home country. This is in accordance with the mechanism described in section 3.4, that a job seeker can receive job information in any occupation where established co-ethnics are employed.

We therefore examine the relationship between the size of the immigrant's network of contacts in a given occupation and their likelihood of being observed in that occupation, using an estimator which allows for a larger set of alternative outcomes, the Alternative-specific conditional logit (Asclogit). The Asclogit has the advantage of allowing for both choice attributes (for example one's network of social contacts in an occupation) and individual attributes in the estimation. Unlike Patel and Vella (2013), our outcome variable consists of 9 occupations in this case. We briefly outline below how this approach compares to the one adopted in Patel and Vella.

Columns 1 to 6 in Table 3.3 illustrate the structure of the data at the point of estimation of the Asclogit, and columns 8 and 9 illustrate the structure of the data at the point of estimation of the LPM. The variable "Age" is individual specific and "Co-ethnics in occupation" is an attribute that varies across occupation and country of birth. While the variable "Occupation chosen" shows the observed occupation of the individual, "Alternative occupation" indicates all possible alternatives they faced. "Whether occupation chosen" takes the value of 1 when "Alternative occupation" is the same as "Occupation chosen" and takes the value of 0 otherwise. This is the dependent variable in the Asclogit. "Popular occupation chosen" is a dummy denoting whether the popular occupation is chosen or not. This is the dependent variable used by Patel and Vella.

For ease of comparison, it is perhaps worth substituting the LPM with a logit estimator. This ought to maintain the comparison without much loss of generality while making this much simpler. The probability of interest in Patel and Vella is:

$$P_o(Y = 1 | o = \text{most popular}) | X \quad (3.5.1)$$

where  $Y$  takes the value of 1 if the individual's occupation corresponds to the most popular occupation amongst their established co-ethnics counterparts, and 0 otherwise. Denoting the positive outcome as  $o_1$  and the negative outcome as  $o_0$ , we have:

$$P_{o_1} = P(Y = o_1 | X, o) = \frac{e^{\beta_{o_1} X}}{\sum_{k \in O} e^{\beta_k X}} \quad (3.5.2)$$

Normalising  $\beta_{o_0}$ , we get  $P_{o_1} = P(Y = o_1 | X, o) = \frac{e^{\beta_{o_1} X}}{1 + e^{\beta_{o_1} X}}$  and

$$\frac{P_{o_1}}{P_{o_0}} = e^{\beta_{o_1} X}, \text{ with } \log\left[\frac{P_{o_1}}{P_{o_0}}\right] = \beta_{o_1} X \quad (3.5.3)$$

Note that the network variable enters the utility obtained from a given outcome as an individual attribute here. As outlined in Cameron and Trivedi (2009), estimating the Asclogit is equivalent to estimating a multinomial logit in the absence of choice attributes. Notwithstanding this specification requirement, assume that one allows the occupation attribute to enter the multinomial logit as an explanatory variable. One can re-structure the data with occupations in order of popularity, and then estimate a pairwise logit while maintaining a given occupation as the base outcome. This would be the same as estimating a multinomial logit with the same base outcome (although at the cost of possibly less efficient coefficient estimates). What is first needed for such methods to be comparable to the one adopted in Patel and Vella is that one chooses the base outcome to be the most popular occupation. Second, the odds between choosing the most popular occupation and any other occupation would have the same response to the size of the network of social contacts in that occupation, regardless of what the occupation is. This key restriction would make our approach collapse into the same approach adopted in Patel and Vella. In the last part of the chapter, we move away from such a restriction and estimate the relationship between the size of the job seeker's network of social contacts in a given occupation and their likelihood of finding employment in that occupation.

Table 3.3: Template of the data before estimation of the Asclogit and the LPM

| <i>ID</i> | <i>Age</i> | <i>Alternative occupation</i> | <i>Whether occupation chosen</i> | <i>Occupation chosen</i> | <i>% of co-ethnics in occupation</i> | <i>Popular occupation</i> | <i>Popular occupation chosen</i> | <i>% of co-ethnics in popular occupation</i> |
|-----------|------------|-------------------------------|----------------------------------|--------------------------|--------------------------------------|---------------------------|----------------------------------|----------------------------------------------|
| 1         | 41         | 1                             | 0                                | 3                        | 11.1                                 | 3                         | 1                                | 24.4                                         |
| 1         | 41         | 2                             | 0                                | 3                        | 11.1                                 |                           |                                  |                                              |
| 1         | 41         | 3                             | 1                                | 3                        | 24.4                                 |                           |                                  |                                              |
| 1         | 41         | 4                             | 0                                | 3                        | 22.2                                 |                           |                                  |                                              |
| 1         | 41         | 5                             | 0                                | 3                        | 11.1                                 |                           |                                  |                                              |
| 1         | 41         | 6                             | 0                                | 3                        | 10.0                                 |                           |                                  |                                              |
| 1         | 41         | 7                             | 0                                | 3                        | 5.0                                  |                           |                                  |                                              |
| 1         | 41         | 8                             | 0                                | 3                        | 3.0                                  |                           |                                  |                                              |
| 1         | 41         | 9                             | 0                                | 3                        | 2.0                                  |                           |                                  |                                              |
| 2         | 26         | 1                             | 0                                | 6                        | 12.5                                 | 2                         | 0                                | 37.5                                         |
| 2         | 26         | 2                             | 0                                | 6                        | 37.5                                 |                           |                                  |                                              |
| 2         | 26         | 3                             | 0                                | 6                        | 12.5                                 |                           |                                  |                                              |
| 2         | 26         | 4                             | 0                                | 6                        | 12.5                                 |                           |                                  |                                              |
| 2         | 26         | 5                             | 0                                | 6                        | 6.3                                  |                           |                                  |                                              |
| 2         | 26         | 6                             | 1                                | 6                        | 18.8                                 |                           |                                  |                                              |
| 2         | 26         | 7                             | 0                                | 6                        | 0.0                                  |                           |                                  |                                              |
| 2         | 26         | 8                             | 0                                | 6                        | 0.0                                  |                           |                                  |                                              |
| 2         | 26         | 9                             | 0                                | 6                        | 0.0                                  |                           |                                  |                                              |
| 3         | 24         | 1                             | 0                                | 9                        | 14.3                                 | 7                         | 0                                | 28.6                                         |
| 3         | 24         | 2                             | 0                                | 9                        | 14.3                                 |                           |                                  |                                              |
| 3         | 24         | 3                             | 0                                | 9                        | 14.3                                 |                           |                                  |                                              |
| 3         | 24         | 4                             | 0                                | 9                        | 14.3                                 |                           |                                  |                                              |
| 3         | 24         | 5                             | 0                                | 9                        | 0.0                                  |                           |                                  |                                              |
| 3         | 24         | 6                             | 0                                | 9                        | 0.0                                  |                           |                                  |                                              |
| 3         | 24         | 7                             | 0                                | 9                        | 28.6                                 |                           |                                  |                                              |
| 3         | 24         | 8                             | 0                                | 9                        | 14.3                                 |                           |                                  |                                              |
| 3         | 24         | 9                             | 1                                | 9                        | 0.0                                  |                           |                                  |                                              |

### 3.6 Econometric analysis

This section starts with testing the hypothesis that the likelihood of obtaining employment through referral in a given occupation increases in the size of one's network of social contact in that occupation – i.e. *hypothesis 1*. We then examine whether the network of social contacts leads immigrants to choosing the most popular occupation amongst fellow co-ethnics already established in their respective employment (i.e. *hypothesis 2*). The section concludes with looking at the relationship between the size of one's network of social contacts in a given occupation and the likelihood of choosing that occupation.

**Table 3.4: Summary statistic of sample – employees with less than a year with employer**

|                                                              | <i>Immigrants by duration of stay in the UK</i> |                      |                  |                     |                   |
|--------------------------------------------------------------|-------------------------------------------------|----------------------|------------------|---------------------|-------------------|
|                                                              | <i>All</i>                                      | <i>Under 2 years</i> | <i>2-5 years</i> | <i>6 - 10 years</i> | <i>11 years +</i> |
| Mean age                                                     | 32.51<br>(0.10)                                 | 29.72<br>(0.12)      | 31.77<br>(0.15)  | 36.01<br>(0.22)     | 45.32<br>(0.28)   |
| Mean age left full-time education                            | 20.42<br>(0.04)                                 | 20.47<br>(0.05)      | 20.66<br>(0.08)  | 20.37<br>(0.14)     | 19.42<br>(0.14)   |
| Proportion of male (in %)                                    | 53.81<br>(0.01)                                 | 54.92<br>(0.01)      | 53.47<br>(0.01)  | 54.00<br>(0.02)     | 48.76<br>(0.02)   |
| Jobs found through referral (in %)                           | 25.11<br>(0.004)                                | 25.41<br>(0.01)      | 24.24<br>(0.01)  | 23.81<br>(0.01)     | 28.02<br>(0.02)   |
| <b><i>Spatial unit of place of work – region</i></b>         |                                                 |                      |                  |                     |                   |
| Mean established co-ethnics in own occupation                | 1.69<br>(0.14)                                  |                      |                  |                     |                   |
| Mean established co-ethnics in employment                    | 10.68<br>(0.11)                                 |                      |                  |                     |                   |
| Mean established employees in own occupation                 | 458.49<br>(1.98)                                |                      |                  |                     |                   |
| <b><i>Spatial unit of place of work – entire country</i></b> |                                                 |                      |                  |                     |                   |
| Mean established co-ethnics in own occupation                | 14.87<br>(0.10)                                 |                      |                  |                     |                   |
| Mean established co-ethnics in other occupations             | 103.02<br>(0.77)                                |                      |                  |                     |                   |
| Mean established employees in own occupation                 | 5541.53<br>(11.11)                              |                      |                  |                     |                   |
| Observations                                                 | 7,885                                           | 3,869                | 2,364            | 924                 | 728               |

*The sample consists of those who started their current job in the last 3 months prior to the survey before 2005, and those who started their current job in the last 12 months prior to the survey after 2005. Standard error in parenthesis.*

Our sample refers to all immigrants who have been with their current employer for at most a year – i.e. those we define as new employee. When testing the first hypothesis, only those who responded to the question of how employment was

obtained are kept in the sample. Summary statistics of our data are reported in Table 3.4. About a quarter of respondents found employment through referrals across all groups of immigrants with different lengths of stay in the UK. More precisely, it is 28 per cent amongst immigrants with over 10 years of stay in the UK, in comparison to about 24 to 25 per cent for other groups of immigrants. As one may expect, recent immigrants are much younger than more established immigrants. Educational attainment levels, as measured by the age of leaving continuous full-time education, are comparable across immigrants with different lengths of stay in the UK. The sample comprises individuals from 113 countries, working across 19 regions (we aggregate Inner London and Outer London areas).

### **3.6.1 The size of networks in an occupation and the likelihood of finding employment in a given occupation through referral**

We empirically test *Hypothesis 1* in this subsection. The hypothesis is that the larger the number of employed members of the network (i.e. pool of the source of job information) in a segment of the labour market, the higher the probability of finding employment through the network relative to other methods of search in that segment of the labour market will be. Hence our measure of the size of the network is the number of established employees in a given occupation. An alternative measure would be the share of within group employment in the occupation. The advantage of using numbers over shares is that the former allows for some of the variation across groups that would be lost when using shares as a measure (a value of 10 per cent, for example, can be the ratio of either 1 and 10, or that of 2 and 20). We particularly examine how this relationship varies with the duration of stay of the immigrant in the UK.

The network measure is constructed at either the regional level or countrywide level depending on the spatial unit we look at. Molly et al (2011) note that inter-regional migration in Great Britain was less than 12 per cent in 2005 when the spatial unit of analysis is at the NUTS2 level – NUTS2s are smaller spatial units than regions of work, there are 36 of them in Great Britain. Furthermore, subdividing England and Wales into 9 main regions, estimates of gross inter-regional migration (i.e. those who changed their region of residence) in mid-2013 ranged between 3.3 and 5.7 per cent (Internal Migration by Local Authorities in

England and Wales, ONS, 2013<sup>15</sup>), confirming that internal migration in the country is quite low. The corresponding figures are likely to be even lower were we to consider changes in region of work instead of region of residence. Our data has 19 regions. In light of this and in order to overcome data limitations, we measure network as an average of established employees over a five year period. Hence, this variable is a moving average.

We initially estimate the following baseline equation:

$$Pr_{gro,t}(job\ obtained\ through\ referral) = \beta_0 + \beta_1 ctrymn_{gro,t}^{estab} * duration + \beta_2 all_{gr,t}^{estab} + \beta_3 birthplace + \beta_4 occupation + \beta_5 region + \beta_6 year + \beta_7 tenure + \varepsilon_{gro,t} \quad (3.6.1)$$

$ctrymn_{gro,t}^{estab}$  measures the number of established fellow co-ethnics in region  $r$ , of the new employee, who are in the same occupation as them.  $all_{gr,t}^{estab}$  denotes the number of all established co-ethnics in employment in the region.  $tenure$  is a dummy which we include in the estimation in order to account for the change in eligibility to respond to the question about how one obtained employment from 2005 onward.

As suggested by Munshi (2003) and Kanas et al (2012), the most immediate source of information for new immigrants is likely to be their co-ethnics in the host country, hence one would expect referral from co-ethnics have a stronger influence amongst new immigrants. At the same time, as an immigrant spends more time in the UK, the available sources of job information to them can be expected to increase. Indirect evidence for this is the literature that shows that immigrants experience better labour market outcomes with duration of stay in the host country (as seen in the literature review section of the previous chapter), as they become accustomed to the host country's labour market. In addition, the pool of contacts of an immigrant may enlarge with duration of stay in the country, such that non-co-ethnics become sources of referrals. Our coefficient of interest,  $\beta_1$ , is therefore computed for groups of new employees with different durations of stay in the UK. This is captured by the interaction term  $ctrymn_{gro,t}^{estab} * duration$  in equation (3.6.1).  $duration$  takes 4 values depending on

---

<sup>15</sup> See:

<http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/re/migration1/internal-migration-by-local-authorities-in-england-and-wales/index.html>

the length of stay of the new employee in the UK: less than 2 years in the UK; 2-5 years in the UK; 6-10 years in the UK; and over 10 years in the UK.

Consistent with the mechanism described in section 4, we have included region of work, occupation and year dummies in the estimation. Regional dummies ought to account of differences in levels of homophily across regions that could arise from, for example, their industrial compositions. Time dummies are included to account for macro shocks. And as we saw earlier in Table 3.2, referral rates differ greatly across occupations, hence the inclusion of occupational dummies in the regression.

We first extend equation (3.6.1) to account for individual characteristics because of our assumption that all individuals are identical within each subgroup in section 3.4. We later include the number of all established employees in occupation  $o$  in the equation. This plays a dual role: on the one hand, it controls for the size of the occupation when the job seeker was looking for employment. On the other hand, job seekers from different countries may be competing for the same job information within a given segment of the labour market given established employees from each country relay job information/provide referrals to their fellow co-ethnics in their respective segments of the labour market. Our extended equation becomes:

$$\begin{aligned} Pr_{gro,t}(\text{job obtained through referral}) = & \beta_0 + \beta_1 \text{ctrym}_{gro,t}^{\text{estab}} * \text{duration} + \\ & \beta_2 \text{all}_{gr,t}^{\text{estab}} + \beta_3 \text{all}_{ro,t}^{\text{estab}} + \beta_4 \text{occupation} + \beta_5 \text{education} + \beta_6 \text{age} + \beta_7 \text{sex} + \\ & \beta_8 \text{birthplace} + \beta_9 \text{region} + \beta_{10} \text{year} + \beta_{11} \text{tenure} + \varepsilon_{gro,t} \end{aligned} \quad (3.6.2)$$

$\text{all}_{ro,t}^{\text{estab}}$  measures the pool of all established employees, including natives, who are in the same occupation as the new employee.

While individuals are more likely to have face to face interactions at the local level, technological advances mean that two individuals living in different regions might be as connected as those living in the same locality. We therefore start with the region of work as the unit of our spatial analysis and then extend it to cover the entire country. We estimate equations (3.6.1) and (3.6.2) using a logit estimator.

## **Results**

Table 3.5 presents our coefficient estimates from equations (3.6.1) and (3.6.2) in 2 sets of columns. The first set contains estimates when the spatial unit of interest is the region of work and the last set contains estimates when the spatial unit of interest is the entire country. We see from column (1) that when the spatial unit of analysis is the region of work, a one-unit increase in the size of the network of social contacts – defined as fellow co-ethnics – in an occupation is associated with an increase in the likelihood of having obtained employment in that occupation through referral by 2.2 per cent amongst new employees with less than 2 years of stay in the UK. The inclusion of individual characteristics such as age, gender and educational attainment hardly affects our coefficient estimates of interest.

Including all established employees – i.e. the size of the occupation when the job seeker was looking for employment – in the estimation as reported in column (3), we now see that a unit increase in the size of the network in an occupation is associated with an increase in the likelihood of obtaining employment through referral in that occupation by 2.7 per cent. This marginal increase relative to the initial estimates could be driven by the fact that all established employees in the occupation include individuals other than fellow co-ethnics of the job seeker, and these individuals would also provide referral to their own compatriots which would thereby lead to some competition.

Moving down column (1), we consider migrants who have been in the UK for longer. We see that the magnitudes of the coefficients of interest drop considerably and lose statistical significance, albeit with no clear pattern. A one unit increase in the size of network in an occupation leads to respectively a 1 per cent, a 1.6 per cent and a 1.9 per cent increase in the likelihood of having obtained employment in that occupation through referral for new employees with 2 to 5 years in the UK, 6 to 10 years in the UK and over 10 years in the UK. Column (3), where estimates from equation (3.6.2) are reported, leads to similar conclusions as column (1). Here, however, the effect of the size of network of social contacts in an occupation on the likelihood of obtaining employment through referral in that occupation is only statistically significant for immigrants with less than 2 years of stay in the UK.

We saw earlier that the shares of jobs found through referral amongst immigrants with lengthier durations of stay in the UK remain similar to the shares of jobs found through referral amongst immigrants with less than 2 years of stay in the UK. As such, these findings suggest fellow co-ethnics established in their own occupations (our definition of the migrant's network) play a more important role in leading new immigrants to choosing the same occupation as members of the network. It is possible that, with time spent in the UK, the networks of social contacts of migrant job seekers no longer predominantly consist of fellow co-ethnics established in their respective occupations; these potentially broaden to include natives and as well as immigrants from other countries<sup>16</sup>.

Columns (5) to (7) of Table 3.5 report estimated coefficients from equations (3.6.1) and (3.6.2) when the spatial unit of interest is the entire country. As seen from column (6), now an increase of one unit in the size of the network in a given occupation is associated with an increase in the likelihood of obtaining employment through referral in that occupation by 0.3 amongst new employees with less than 2 years of stay in the UK (compare to 2.7 per cent when the spatial unit of analysis is the region of work). Here too, we see that the magnitudes of the coefficients of interest are generally much smaller and statistically insignificant at conventional levels for new employees with longer lengths of stay in the UK, except amongst immigrants with over 10 years of stay in the UK. The magnitude of the coefficient of interest for this group of immigrants is over twice that of the magnitude of the coefficient of interest for immigrants with less than 2 years of stay in the UK, albeit statistically insignificant.

The negative coefficients on the size of all employed individuals in the subgroup at the time when the job seeker was looking for employment in the first three columns of results, regardless of duration of stay in the UK, possibly suggest members of the network in each segment of the labour market are more likely to provide job information/referrals in their own segments of the labour market, rather than other segments of the labour market. Coefficient estimates for the size of all established employees in the occupation variables are all zero, which could be because the competition effect and the occupation size effect cancel each other out here. The likelihood of obtaining employment through the

---

<sup>16</sup> For instance, in a 2012 article by the Economist, it is suggested that 8.8 per cent of marriages in Britain include a foreign-born spouse: see <http://www.economist.com/blogs/feastandfamine/2012/07/mixed-marriages>

network reduces with educational attainment, which is consistent with findings from other studies (see Wahba and Zenou (2005), for example).

We re-estimate our coefficients of interest with immigrants from the A8 countries (i.e. the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia) observed from 2004 onward excluded from the sample, given that since the accession of the A8 countries to the EU, there has been a sharp increase in the stock of immigrants from these countries in the UK. Additionally, empirical evidence suggests a high propensity of finding employment through referrals amongst these individuals (see Sumption (2009) on immigrants from Poland). This allows us to examine the extent to which this group of immigrants could be driving the results. Coefficient estimates from equation (3.6.2) with this restricted sample are reported in column (1) and column (2) of Table 3.6 for when the spatial unit of interest is respectively the region and the entire country. The new coefficient estimates are now 0.033 (0.005) – with a 95% CI of [0.001 0.041] – as opposed to 0.027 (0.003) when the spatial unit of analysis is the region of work (the entire country), indicating only a slight increase in the magnitude of the initial estimates.

Another concern is the number of small values in our key explanatory variable of interest. As an additional robustness check, we restrict the sample to observations with network size between the 10<sup>th</sup> and the 90<sup>th</sup> percentile, and re-estimate equation (6.2). Coefficient estimates from this trimmed sample are reported in columns (3) and (4) of Table 3.6. We see that the initial estimates obtained with the full sample remain robust, particularly when the spatial unit is the region of work. The new estimates are now 0.029 (0.001) as opposed to 0.027 (0.003) when the spatial unit of analysis is the region of work (the entire country), indicating again only a slight increase from our initial estimates. Result from this exercise partly alleviate concerns for a number of individuals in our sample having zero as the value of their network size, too. Nevertheless, we re-estimate equation (3.6.2) with a sample where we drop observations with a network size value of zero. These observations represent slightly over 10 per cent of our sample. We find that the resulting estimates from this exercise are the same as those reported in Table 3.5, our main results. See Table 3.7 for the new estimates.

Our final robustness check consists of estimating equation (3.6.2) with standard errors clustered by country birth and region of work, unlike before when we only

clustered standard errors by country of birth. To circumvent the lack of readily available software that computes marginal effects after estimating a logit with a two way clustered standard errors, we estimate a LPM instead. In the interest of clarity, we estimate equation (3.6.2) first with standard error clustered by country of birth, and then again with standard errors clustered by country birth and region of work. As can be seen from Table 3.8, the resulting estimates show that our estimate of interest remain robust to clustering standard errors by country of birth and region of work when the spatial unit of analysis is region of work, but we no longer have statistical significance when the spatial unit of analysis is the entire country. With a little computation, it is straight forward to compare these new estimates to those obtained from using a logit. For example, new employees with less than two years of stay in the UK have a 0.25 probability of finding employment through referrals in the sample. Table 3.8 reports that a unit increase in the network size is associated with a 0.008 increase in likelihood of obtaining employment through referral in a particular occupation for this group of immigrants. This corresponds to a 3.2 per cent increase, compared with 2.7 per cent from our main results.

Given that our outcome variable is whether employment was obtained through referral or not in equations (3.6.1) and (3.6.2), this does not shed light as to other means through which one's network of social contacts could influence the occupational outcomes of the job seeker. This is because established employees within the network could for example promote their respective occupations to the job seeker. The job seeker could also potentially receive guidance from the network such as which employment agency provides good job search services in their respective occupations, newspapers where they may find information about job vacancies, etc. A paper which deals with such shortcomings is Patel and Vella (2013). They examine whether a new immigrant chooses the most popular occupation amongst established fellow immigrants in the context of the USA. We therefore proceed in a similar vein as Patel and Vella (2013) to examine whether newly employed immigrants are likely to be in the most popular occupation amongst established co-ethnics.

**Table 3.5: Likelihood of finding employment in an occupation through referral and the size of the network in the occupation: semi-elasticities from equations (3.6.1) and (3.6.2), reported across new employees with different durations of stay in the UK**

|                                                                    | <i>Region of work</i> |                       |                       | <i>Entire country</i> |                       |                       |
|--------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                                                                    | (M1)                  | (M2)                  | (M3)                  | (M1)                  | (M2)                  | (M3)                  |
| <b><i>Established co-ethnics in own occupation, effects on</i></b> |                       |                       |                       |                       |                       |                       |
| Migrants with < 2 years in the UK                                  | 0.022**<br>(2.132)    | 0.021**<br>(2.095)    | 0.027***<br>(2.670)   | 0.003**<br>(2.419)    | 0.003**<br>(2.398)    | 0.003**<br>(2.071)    |
| Migrants with 2-5 years in the UK                                  | 0.009<br>(0.757)      | 0.008<br>(0.690)      | 0.009<br>(0.711)      | -0.001<br>(-0.700)    | -0.001<br>(-0.695)    | -0.001<br>(-0.777)    |
| Migrants with 6-10 years in the UK                                 | 0.016<br>(1.215)      | 0.015<br>(1.116)      | 0.002<br>(0.117)      | -0.003<br>(-0.632)    | -0.003<br>(-0.658)    | -0.004<br>(-1.043)    |
| Migrants with 10+ years in the UK                                  | 0.019*<br>(1.889)     | 0.018*<br>(1.764)     | 0.016<br>(1.559)      | 0.008<br>(1.610)      | 0.008<br>(1.640)      | 0.007<br>(1.543)      |
| <b><i>Established co-ethnics in employment, effects on</i></b>     |                       |                       |                       |                       |                       |                       |
| Migrants with < 2 years in the UK                                  | -0.003**<br>(-2.435)  | -0.002**<br>(-2.471)  | -0.002**<br>(-2.157)  | 0.001<br>(0.947)      | 0.001<br>(0.987)      | 0.001<br>(0.982)      |
| Migrants with 2-5 years in the UK                                  | -0.003**<br>(-2.427)  | -0.003**<br>(-2.464)  | -0.002**<br>(-2.150)  | 0.001<br>(0.947)      | 0.001<br>(0.986)      | 0.001<br>(0.981)      |
| Migrants with 6-10 years in the UK                                 | -0.003**<br>(-2.483)  | -0.003**<br>(-2.518)  | -0.002**<br>(-2.196)  | 0.001<br>(0.932)      | 0.001<br>(0.971)      | 0.001<br>(0.968)      |
| Migrants with 10+ years in the UK                                  | -0.002**<br>(-2.454)  | -0.002**<br>(-2.497)  | -0.002**<br>(-2.169)  | 0.001<br>(0.936)      | 0.001<br>(0.977)      | 0.001<br>(0.975)      |
| <b><i>All established employees in occupation, effects on</i></b>  |                       |                       |                       |                       |                       |                       |
| Migrants with < 2 years in the UK                                  |                       |                       | -0.0002<br>(-1.029)   |                       |                       | -0.000003<br>(-0.062) |
| Migrants with 2-5 years in the UK                                  |                       |                       | -0.00005<br>(-0.274)  |                       |                       | 0.00001<br>(0.174)    |
| Migrants with 6-10 years in the UK                                 |                       |                       | 0.001***<br>(2.577)   |                       |                       | 0.0001***<br>(2.857)  |
| Migrants with 10+ years in the UK                                  |                       |                       | 0.0002<br>(0.749)     |                       |                       | 0.0001***<br>(3.037)  |
| <b><i>Individual characteristics</i></b>                           |                       |                       |                       |                       |                       |                       |
| Age                                                                |                       | -0.001<br>(-0.435)    | -0.001<br>(-0.345)    |                       | -0.001<br>(-0.605)    | -0.001<br>(-0.457)    |
| Education                                                          |                       | -0.023***<br>(-3.968) | -0.023***<br>(-4.029) |                       | -0.022***<br>(-3.891) | -0.023***<br>(-3.976) |
| Male                                                               |                       | 0.040<br>(0.641)      | 0.043<br>(0.691)      |                       | 0.041<br>(0.701)      | 0.039<br>(0.658)      |

Standard errors clustered by country of origin. Significant levels: \*.10, \*\*0.05, \*\*\*0.01. t-statistics in parenthesis. Estimates for each duration of stay obtained by interacting the network variable with the duration of stay variable.

**Table 3.5: continues...**

|                                  |                      |                     |                     |                      |                     |                     |
|----------------------------------|----------------------|---------------------|---------------------|----------------------|---------------------|---------------------|
| <b>Occupation</b>                |                      |                     |                     |                      |                     |                     |
| Professional                     | -0.417**<br>(-2.186) | -0.361*<br>(-1.902) | -0.326<br>(-1.488)  | -0.416**<br>(-2.183) | -0.360*<br>(-1.897) | -0.225<br>(-1.111)  |
| Associate professional           | -0.041<br>(-0.279)   | -0.023<br>(-0.157)  | 0.016<br>(0.092)    | -0.045<br>(-0.321)   | -0.026<br>(-0.183)  | 0.082<br>(0.497)    |
| Clerical and secretarial         | -0.069<br>(-0.439)   | -0.070<br>(-0.461)  | -0.035<br>(-0.197)  | -0.094<br>(-0.639)   | -0.093<br>(-0.640)  | -0.007<br>(-0.040)  |
| Skill trade or craft             | 0.698***<br>(4.974)  | 0.632***<br>(4.409) | 0.680***<br>(3.463) | 0.668***<br>(5.106)  | 0.605***<br>(4.516) | 0.753***<br>(4.504) |
| Personal service                 | 0.256**<br>(1.964)   | 0.238*<br>(1.888)   | 0.287<br>(1.509)    | 0.232*<br>(1.924)    | 0.217*<br>(1.849)   | 0.415**<br>(2.096)  |
| Sales                            | 0.398**<br>(2.478)   | 0.377**<br>(2.419)  | 0.415*<br>(1.811)   | 0.366**<br>(2.568)   | 0.348**<br>(2.511)  | 0.559**<br>(2.452)  |
| Machine operative                | 0.378**<br>(2.371)   | 0.315*<br>(1.942)   | 0.356<br>(1.519)    | 0.347**<br>(2.335)   | 0.286*<br>(1.894)   | 0.485**<br>(2.075)  |
| Elementary                       | 0.642***<br>(4.704)  | 0.593***<br>(4.439) | 0.637***<br>(3.237) | 0.615***<br>(4.783)  | 0.569***<br>(4.486) | 0.745***<br>(4.100) |
| Less than 3 months with employer | 0.004<br>(0.067)     | 0.006<br>(0.094)    | 0.007<br>(0.108)    | 0.008<br>(0.136)     | 0.009<br>(0.156)    | 0.008<br>(0.138)    |
| Year of survey dummies           | Yes                  | Yes                 | Yes                 | Yes                  | Yes                 | Yes                 |
| Region of work dummies           | Yes                  | Yes                 | Yes                 | Yes                  | Yes                 | Yes                 |
| Birthplace dummies               | Yes                  | Yes                 | Yes                 | Yes                  | Yes                 | Yes                 |
| <b>Observations</b>              | <b>7,885</b>         | <b>7,885</b>        | <b>7,885</b>        | <b>7,885</b>         | <b>7,885</b>        | <b>7,885</b>        |

Standard errors clustered by country of origin. Significant levels: \*.10, \*\*.05, \*\*\*0.01. t-statistics in parenthesis. Estimates for each duration of stay obtained by interacting the network variable with the duration of stay variable.

**Table 3.6: Likelihood of finding employment in an occupation through referral and the size of the network in the occupation: semi-elasticities from equation (3.6.2), reported across new employees with different durations of stay in the UK. Restricted samples**

|                                                                    | <i>Excludes immigrants<br/>from A8 countries</i> |                      | <i>Trimmed sample</i> |                      |
|--------------------------------------------------------------------|--------------------------------------------------|----------------------|-----------------------|----------------------|
|                                                                    | (1)                                              | (2)                  | (3)                   | (4)                  |
| <b><i>Established co-ethnics in own occupation, effects on</i></b> |                                                  |                      |                       |                      |
| Migrants with < 2 years in the UK                                  | 0.033**<br>(2.386)                               | 0.005<br>(1.406)     | 0.029***<br>(2.955)   | 0.001<br>(0.529)     |
| Migrants with 2-5 years in the UK                                  | 0.015<br>(1.216)                                 | -0.002<br>(-0.589)   | 0.008<br>(0.645)      | -0.001<br>(-0.541)   |
| Migrants with 6-10 years in the UK                                 | 0.012<br>(0.911)                                 | 0.000<br>(0.039)     | 0.003<br>(0.177)      | -0.006<br>(-1.019)   |
| Migrants with 10+ years in the UK                                  | 0.018*<br>(1.762)                                | 0.009<br>(1.377)     | 0.020**<br>(1.991)    | 0.009**<br>(2.206)   |
| <b><i>Established co-ethnics in employment, effects on</i></b>     |                                                  |                      |                       |                      |
| Migrants with < 2 years in the UK                                  | -0.003***<br>(-2.685)                            | -0.001*<br>(-1.735)  | -0.002*<br>(-1.865)   | 0.001<br>(0.964)     |
| Migrants with 2-5 years in the UK                                  | -0.003***<br>(-2.657)                            | -0.001*<br>(-1.729)  | -0.002*<br>(-1.865)   | 0.001<br>(0.965)     |
| Migrants with 6-10 years in the UK                                 | -0.003***<br>(-2.659)                            | -0.001*<br>(-1.739)  | -0.002*<br>(-1.902)   | 0.001<br>(0.951)     |
| Migrants with 10+ years in the UK                                  | -0.003***<br>(-2.681)                            | -0.001*<br>(-1.751)  | -0.002*<br>(-1.877)   | 0.001<br>(0.958)     |
| <b><i>All established employees in occupation, effects on</i></b>  |                                                  |                      |                       |                      |
| Migrants with < 2 years in the UK                                  | -0.0002<br>(-0.817)                              | -0.00001<br>(-0.148) | -0.0002<br>(-0.977)   | -0.00004<br>(-0.840) |
| Migrants with 2-5 years in the UK                                  | 0.00001<br>(0.056)                               | 0.00003<br>(0.728)   | 0.00005<br>(0.290)    | -0.00003<br>(-0.698) |
| Migrants with 6-10 years in the UK                                 | 0.001**<br>(2.163)                               | 0.0001***<br>(2.721) | 0.0007***<br>(2.661)  | 0.0001**<br>(2.222)  |
| Migrants with 10+ years in the UK                                  | 0.0002<br>(1.074)                                | 0.0001***<br>(3.212) | 0.0002<br>(0.730)     | 0.0001**<br>(2.476)  |
| Observations                                                       | 6,107                                            | 6,107                | 6,579                 | 6,796                |

Standard errors clustered by country of origin. Significant levels: \*.10, \*\*.05, \*\*\*0.01. t-statistics in parenthesis. Estimates for each duration of stay obtained by interacting the network variable with the duration of stay variable. Controls variables remain the same as those in equation (3.6.2).

**Table 3.7: Likelihood of finding employment in an occupation through referral and the size of the network in the occupation: semi-elasticities from equation (3.6.2), reported across new employees with different durations of stay in the UK. Sample excludes those with a network size of zero**

|                                                             | Region of work      | Entire county      |
|-------------------------------------------------------------|---------------------|--------------------|
| <b>Established co-ethnics in own occupation, effects on</b> |                     |                    |
| Migrants with < 2 years in the UK                           | 0.027***<br>(2.622) | 0.003**<br>(2.017) |
| Migrants with 2-5 years in the UK                           | 0.009<br>(0.751)    | -0.001<br>(-0.804) |
| Migrants with 6-10 years in the UK                          | 0.003<br>(0.171)    | -0.004<br>(-1.035) |
| Migrants with 10+ years in the UK                           | 0.016<br>(1.534)    | 0.007<br>(1.509)   |
| N                                                           | 7,167               | 7,863              |

Controls variables remain the same as those in equation (3.6.2). Standard errors clustered by country of birth. t-statistics in parentheses.

**Table 3.8: Robustness check with standard errors clustered by (1) country of birth; (2) country of birth and region of work. Average marginal effects from equation (3.6.2) using LPM, reported across new employees with different durations of stay in the UK.**

|                                                             | Region of work      |                    | Entire county       |                     |
|-------------------------------------------------------------|---------------------|--------------------|---------------------|---------------------|
|                                                             | (1)                 | (2)                | (1)                 | (2)                 |
| <b>Established co-ethnics in own occupation, effects on</b> |                     |                    |                     |                     |
| Migrants with < 2 years in the UK                           | 0.008***<br>(3.314) | 0.008**<br>(2.172) | 0.001***<br>(2.654) | 0.001<br>(1.280)    |
| Migrants with 2-5 years in the UK                           | 0.003<br>(1.271)    | 0.003<br>(0.902)   | -0.0001<br>(-0.210) | -0.0001<br>(-0.094) |
| Migrants with 6-10 years in the UK                          | 0.001<br>(0.309)    | 0.001<br>(0.223)   | -0.001<br>(-0.853)  | -0.001<br>(-0.536)  |
| Migrants with 10+ years in the UK                           | 0.005*<br>(1.842)   | 0.005<br>(1.024)   | 0.002<br>(1.585)    | 0.002<br>(0.993)    |
| N                                                           | 7,885               | 7,885              | 7,885               | 7,885               |

Controls variables remain the same as those in equation (3.6.2).. Standard errors clustered by country of birth in (1) and by country of birth and region of work in (2). t-statistics in parentheses.

### 3.6.2 Replicating Patel and Vella (2013)

This section tests *Hypothesis 2*, that is, referrals and job information from the network may lead to employment in the most popular occupation amongst fellow co-ethnics. The estimation strategy adopted here follows Patel and Vella (2013) where they examine whether social networks of contacts cause immigrants to

select the most popular occupation amongst fellow established co-ethnics in the context of the USA. We first estimate the following baseline equation:

$$Pr [(inpop_{ogr,t} = 1)] = \alpha + \gamma_1 ctrymn_{ogr,t}^{inpop} * duration + \varepsilon_{iogr,t} \quad (3.6.3)$$

$inpop_{ogr,t} = 1$  if the new employee from country  $g$ , working in region  $r$ , chooses the most popular occupation  $o$  amongst their more established fellow co-ethnics.

Consistent with Patel and Vella (2013),  $ctrymn_{ogr,t}^{inpop}$  measures the share of established immigrants from country  $g$ , in region  $r$ , who work in the most popular occupation  $o$  amongst all co-ethnics established in their respective employment. This is because the key question of interest here is to examine the strength of the network in channelling its job seeker members toward the predominant occupation amongst those who are in the position to provide referrals and job information. As such, the specification requires that network is measured in shares rather than in absolute numbers.

We later extend equation (3.6.3) to include country of birth dummies, region of work dummies and occupation dummies to respectively account for the propensity of immigrants from a particular country to choose a given occupation, attributes particular to a region which may lead individuals to choosing a given occupation, and occupations which may attract particular groups of immigrants. Individual characteristics are also accounted for here. As in Patel and Vella, we include in the extended version of equation (3.6.3) the share of workers in region  $r$  employed in occupation  $o$  to account for demand in occupation  $o$ . These variables could be considered as determinants of the deterministic part of equation (3.4.6) which does not depend on the size of network of social contacts. This is therefore our preferred specification as it accounts for the correlated and the exogenous effects:

$$\begin{aligned} Pr [(inpop_{ogr,t} = 1)] = & \alpha + \gamma_1 ctrymn_{ogr,t}^{inpop} * duration + \gamma_2 Share_{or,t} + \\ & \gamma_3 natives_{or,t} + \gamma_4 education + \gamma_5 age + \gamma_6 sex + \gamma_7 birthplace + \\ & \gamma_8 occupation + \gamma_9 region + \gamma_{10} year + \varepsilon_{iogr,t} \end{aligned} \quad (3.6.4)$$

$Share_{or,t}$  denotes the share of workers in region  $r$  employed in occupation  $o$ , which we use as a proxy for total demand as do Patel and Vella.  $natives_{or,t}$

measures the share of the natives born in the most popular occupation. In both equations (3.6.3) and (3.6.4), our coefficient of interest remains  $\gamma_1$ , and we examine how it varies with duration of stay in the UK. As we no longer use information on whether employment was obtained through referral, the sample now includes all those who have been with their current employer for at most one year and not just those eligible to answer the question on how employment was obtained as in section (3.6.2). Estimates of  $\gamma_1$  are reported in Table 3.9. As in Patel and Vella (2013), both equation (3.6.3) and equation (3.6.4) are estimated using a LPM.

### **Results:**

The left half of Table 3.9 reports estimates of  $\gamma_1$  when the spatial of analysis is the region of work and the right half reports estimates when the spatial unit of analysis is the entire country. For each set of estimates, the first columns contain estimates from our baseline specification and the remaining columns contain estimates from our preferred specification.

Estimates from the baseline specification show a positive relationship between the size of the network of established (employee) co-ethnics in the most popular occupation and the likelihood that a newly employed immigrant chooses that occupation. This relationship appears stronger in the initial years of immigrants in the UK and weakens with time spent in the country. Amongst very recent immigrants, a percentage point increase in the size of the network is associated with a 0.60 percentage point increase in the likelihood of being in most popular occupation amongst established fellow co-ethnics in their respective occupations. The corresponding elasticities are respectively 0.55, 0.30 and 0.40 amongst new employees with 2 to 5 years in the UK, 6 to 10 years in the UK and 10+ years of stay in the UK.

The magnitudes of the estimates of interest drop considerably when we control for the country of birth and other characteristics in the estimation – i.e. our preferred specification. Here, we see that a 1 percentage point increase in the size of established co-ethnics in the most popular occupation increases the likelihood of the newly employed being in the most popular occupation by respectively 0.23 percentage point and 0.25 percentage point amongst immigrants with less than 2

years and those with 2 to 5 years of stay in the UK – these two estimates are not statistically different. The corresponding effects of interest are zero for immigrants with 6 to 10 years and those with over 10 years of stay in the UK. When the spatial unit of analysis is the entire country, the size of fellow established co-ethnics in the most popular occupation only affects the occupational outcomes of immigrants with less than 2 years of stay in the UK, as shown in the last column of Table 3.9. We see that a 1 percentage point increase in the size of established co-ethnics in the most popular occupation increases the likelihood of the newly employed being in the most popular occupation by 0.35 percentage point. The corresponding relationships for other groups of immigrants are all zero. These results corroborate our earlier findings that the occupational composition of the network looked at here, i.e. fellow co-ethnics, has a reduced effect on the occupational outcomes of job seekers who are more established in the host country.

As we noted earlier, the use of network of social contacts might be more prevalent in some occupations than others. Equations (3.6.3) and (3.6.4) do not however allow one to examine occupations where one's network of social contacts might affect their occupational outcome the most. This is because only two outcomes are possible: a job seeker either finds employment in the most popular occupation or not. Nevertheless, the job seeker can receive job information on any occupation where fellow established co-ethnics are employed. As such, social networks of contact may vary across occupations but would not vary across immigrants from the same home country. We therefore look at a broader set of possible outcomes, based on 9 1-digit occupation codes, and evaluate the relationship between the size of one's network of social contacts in a given occupation and the likelihood that they choose that occupation.

**Table 3.9: Likelihood that a new employee chose the most popular occupation amongst established co-ethnics – average marginal effects from equations (3.6.3) and (3.6.4), reported across new employees with different durations of stay in the UK**

|                                                                            | Region of work      |                       | Entire country      |                       |
|----------------------------------------------------------------------------|---------------------|-----------------------|---------------------|-----------------------|
|                                                                            | (1)                 | (2)                   | (3)                 | (4)                   |
| <b>Share co-ethnics in popular occupation, average marginal effects on</b> |                     |                       |                     |                       |
| Migrants with less than 2 years in the UK                                  | 0.604***<br>(5.787) | 0.234***<br>(2.587)   | 0.996***<br>(4.241) | 0.353***<br>(2.630)   |
| Migrants with 2 - 5 years in the UK                                        | 0.550***<br>(4.954) | 0.257***<br>(2.833)   | 0.528***<br>(2.803) | -0.042<br>(-0.330)    |
| Migrants with 6 - 10 years in the UK                                       | 0.299**<br>(2.561)  | 0.040<br>(0.366)      | 0.472**<br>(2.414)  | 0.096<br>(0.505)      |
| Migrants with over 10 years in the UK                                      | 0.406***<br>(3.076) | 0.070<br>(0.501)      | 0.369**<br>(2.057)  | -0.099<br>(-0.441)    |
| <b>Controls</b>                                                            |                     |                       |                     |                       |
| Share of natives in occupation                                             |                     | -0.659***<br>(-4.686) |                     | -2.575***<br>(-7.521) |
| Share of employees in occupation                                           |                     | 0.116<br>(0.211)      |                     | -1.889<br>(-1.476)    |
| Education                                                                  |                     | -0.002<br>(-1.008)    |                     | -0.005***<br>(-3.484) |
| Female                                                                     |                     | 0.010<br>(1.001)      |                     | 0.008<br>(1.007)      |
| Age                                                                        |                     | -0.000<br>(-0.372)    |                     | -0.000<br>(-0.400)    |
| Occupation dummies                                                         | No                  | Yes                   | No                  | Yes                   |
| Region dummies                                                             | No                  | Yes                   | No                  | Yes                   |
| Country of birth dummies                                                   | No                  | Yes                   | No                  | Yes                   |
| Year dummy                                                                 | No                  | Yes                   | No                  | Yes                   |
| Observations                                                               | 8,118               | 8,118                 | 10,329              | 10,322                |
| Adj. R2                                                                    | 0.025               | 0.179                 | 0.021               | 0.269                 |

Standard errors clustered by country of origin. Significant levels: \*.10, \*\*.05, \*\*\*0.01. t-statistics in parenthesis.

### 3.6.3 Extending Patel and Vella's approach

Here, we use a multinomial estimator, the alternative specific logit (Asclgit), which allows us to estimate the relationship between the likelihood of an immigrant being observed in one of the 9 occupations and the size of their social network of contacts in this occupation. The Asclgit has the advantage of allowing for both choice attributes and individual attributes in the estimation. The estimated equation is the following, which also derives from equation (3.4.6):

$$\begin{aligned}
 Pr [occupation_{org,t} = o] = & \alpha + \gamma_1 ctry_{mn_{org,t}} * duration + \gamma_2 natives_{or,t} + \\
 & \gamma_3 Share_{or,t} + \beta_{j1} birthplace + \beta_{j2} education + \beta_{j3} age + \beta_{j4} sex + \beta_{j5} region + \\
 & \beta_{j1} birthplace + \beta_{j2} education + \beta_{j3} age + \beta_{j4} sex + \beta_{j5} region + \\
 & \beta_{j6} year + \varepsilon_{ogr,t}
 \end{aligned} \tag{3.6.5}$$

Our coefficient of interest remains  $\gamma_1$ . We estimate equation (3.6.5) separately with samples respectively consisting of immigrants with less than 2 years of stay in the UK and those with more than 10 years of stay in the UK, because obtaining coefficient estimates of interaction terms from the Asclogit is not straightforward.

The network variable is defined as the number of fellow co-ethnics already established in their own occupation, as in equations (3.6.1) and (3.6.2). Similarly to section (3.6.1), this choice is motivated by the increase in variation it allows us to have in the data. Unlike in equations (3.6.3) and (3.6.4), now each individual is assigned both co-ethnics in their own occupation as well as co-ethnics in other occupations. Hence the network variable enters the model as a choice attribute. This is consistent with the fact that the job seeker may receive information on any occupation where fellow established co-ethnics are employed. *natives<sub>org</sub>* and *Share<sub>or</sub>* also enter the estimation as occupation attributes. Here too, we include in equation (3.6.5) all employees already established in their respective occupations as we previously did in equation (3.6.1).

### **Results:**

Table 3.10 presents marginal effects of interest from equation (3.6.5), computed at the mean (MEM). We see that the relationship between a migrant's network of social contacts in an occupation and the likelihood of choosing that occupation varies across occupations. For example, amongst immigrants with less than 2 years of stay in the UK, the magnitude of the estimate for professional occupation is 0.013 compared to 0.023 for elementary occupation. These differences are also apparent when one views the estimates effects in term of percentage changes to base probabilities of being in given occupations: while a unit increase in the size of the network is associated with over 10 per cent increase in the likelihood of choosing a managerial or a craft and skilled trade occupation, a similar increase is associated with less than an 8 per cent increase in the likelihood of choosing an elementary occupation. These observations would therefore suggest that the restriction that the effects of the migrant's network of social contacts on occupational outcomes are equal across occupations, as is the case in equations (3.6.2) and (3.6.4), may not be supported. The advantage of

using the Asclogit is that we are able to relax this restriction, thereby allowing us to capture these differences.

Looking at immigrants who have been in the UK for over 10 years, we observe similar patterns as we did for immigrants with less than 2 years of stay in the UK. Also, consistent with our previous results, networks of social contacts appear to affect the occupational outcomes of immigrants with over 10 years in the UK less than they do for immigrants with less than 2 years of stay in the UK. For instance, while a unit increase in the size of the network is associated with over a 10 per cent increase in the likelihood of choosing a managerial or a clerical and secretarial occupation for new immigrants, a similar increase in the size of the network is associated with only about an 8 per cent increase in the likelihood of choosing one of these occupations for immigrants with over 10 years of stay in the UK.

We further compute average marginal effects of the size of migrant's network of social contacts on the likelihood of being observed in a given occupation, using the computation suggested in Cameron and Trivedi (2009). These estimates corroborate the findings from Tables 3.10 – the estimates in question are reported in Table 3.13 in the appendix.

Table 3.10: Likelihood of choosing a given occupation – marginal effects computed at the mean (MEM)

|                                            | Immigrants with less than 2 years in the UK |               |         |          |         | Immigrants with over 10 years in the UK |               |       |          |       |
|--------------------------------------------|---------------------------------------------|---------------|---------|----------|---------|-----------------------------------------|---------------|-------|----------|-------|
|                                            | Probability in data                         | (1)           |         | (2)      |         | Probability in data                     | (3)           |       | (4)      |       |
|                                            |                                             | dp/dx         | z-value | dp/dx    | z-value |                                         | dp/dx         | z     | dp/dx    | z     |
| <b>Managerial occupation</b>               | 0.061                                       |               |         |          |         | 0.111                                   |               |       |          |       |
| Number of established co-ethnics           |                                             | <b>0.0063</b> | 4.32    | 0.0064   | 4.32    |                                         | <b>0.0092</b> | 3.79  | 0.0093   | 3.84  |
| Number of all established employees        |                                             |               |         | -0.00004 | -1.73   |                                         |               |       | -0.00003 | -0.45 |
| Share of natives                           |                                             | -0.4076       | -4.21   | -0.4016  | -4.19   |                                         | -0.6218       | -3.20 | -0.6221  | -3.21 |
| Share of total employment                  |                                             | 0.6863        | 2.74    | 1.0470   | 3.46    |                                         | 1.0884        | 1.59  | 1.3391   | 1.75  |
| <b>Professional occupation</b>             | 0.132                                       |               |         |          |         | 0.210                                   |               |       |          |       |
| Number of established co-ethnics           |                                             | <b>0.0125</b> | 4.54    | 0.0127   | 4.56    |                                         | <b>0.0156</b> | 3.56  | 0.0157   | 3.61  |
| Number of all established employees        |                                             |               |         | -0.0001  | -1.81   |                                         |               |       | -0.00005 | -0.45 |
| Share of natives                           |                                             | -0.8115       | -4.42   | -0.7987  | -4.38   |                                         | -1.0508       | -3.28 | -1.0498  | -3.28 |
| Share of total employment                  |                                             | 1.3663        | 2.75    | 2.0822   | 3.69    |                                         | 1.8393        | 1.53  | 2.2596   | 1.72  |
| <b>Associate prof. occupation</b>          | 0.141                                       |               |         |          |         | 0.140                                   |               |       |          |       |
| Number of established co-ethnics           |                                             | <b>0.0132</b> | 5.06    | 0.0134   | 5.09    |                                         | <b>0.0113</b> | 3.90  | 0.0114   | 3.93  |
| Number of all established employees        |                                             |               |         | -0.0001  | -1.84   |                                         |               |       | -0.00003 | -0.44 |
| Share of natives                           |                                             | -0.8559       | -5.64   | -0.8429  | -5.54   |                                         | -0.7614       | -2.93 | -0.7603  | -2.92 |
| Share of total employment                  |                                             | 1.4410        | 2.52    | 2.1974   | 3.44    |                                         | 1.3327        | 1.57  | 1.6364   | 1.66  |
| <b>Clerical and secretarial occupation</b> | 0.069                                       |               |         |          |         | 0.172                                   |               |       |          |       |
| Number of established co-ethnics           |                                             | <b>0.0070</b> | 4.83    | 0.0071   | 4.86    |                                         | <b>0.0134</b> | 3.93  | 0.0135   | 3.96  |
| Number of all established employees        |                                             |               |         | -0.00005 | -1.79   |                                         |               |       | -0.00004 | -0.45 |
| Share of natives                           |                                             | -0.4547       | -6.32   | -0.4471  | -6.20   |                                         | -0.9018       | -2.82 | -0.9009  | -2.81 |
| Share of total employment                  |                                             | 0.7656        | 2.42    | 1.1655   | 3.17    |                                         | 1.5785        | 1.53  | 1.9391   | 1.64  |
| <b>Craft and skilled trade occupation</b>  | 0.051                                       |               |         |          |         | 0.043                                   |               |       |          |       |
| Number of established co-ethnics           |                                             | <b>0.0052</b> | 6.82    | 0.0053   | 6.96    |                                         | <b>0.0039</b> | 3.11  | 0.0039   | 3.15  |
| Number of all established employees        |                                             |               |         | -0.00004 | -1.95   |                                         |               |       | -0.00001 | -0.46 |
| Share of natives                           |                                             | -0.3399       | -5.77   | -0.3335  | -5.61   |                                         | -0.2625       | -2.55 | -0.2620  | -2.55 |
| Share of total employment                  |                                             | 0.5722        | 2.27    | 0.8693   | 3.23    |                                         | 0.4594        | 1.41  | 0.5640   | 1.60  |
| <b>Other controls</b>                      |                                             |               |         |          |         |                                         |               |       |          |       |
| Education                                  |                                             | Yes           |         | Yes      |         |                                         | Yes           |       | Yes      |       |
| Male                                       |                                             | Yes           |         | Yes      |         |                                         | Yes           |       | Yes      |       |
| Age                                        |                                             | Yes           |         | Yes      |         |                                         | Yes           |       | Yes      |       |
| Year                                       |                                             | Yes           |         | Yes      |         |                                         | Yes           |       | Yes      |       |

Table 3.10 continues....

|                                     | Immigrants with less than 2 years in the UK |               |         |          |         | Immigrants with over 10 years in the UK |               |         |          |         |
|-------------------------------------|---------------------------------------------|---------------|---------|----------|---------|-----------------------------------------|---------------|---------|----------|---------|
|                                     | Probability in data                         | (M1)          |         | (M2)     |         | Probability in data                     | (M1)          |         | (M2)     |         |
|                                     |                                             | dydx          | z-value | Dydx     | z-value |                                         | dydx          | z-value | dydx     | z-value |
| <b>Personal service occupation</b>  | 0.127                                       |               |         |          |         | 0.080                                   |               |         |          |         |
| Number of established co-ethnics    |                                             | <b>0.0121</b> | 5.11    | 0.0122   | 5.21    |                                         | <b>0.0069</b> | 4.02    | 0.0070   | 4.07    |
| Number of all established employees |                                             |               |         | -0.0001  | -1.94   |                                         |               |         | -0.00002 | -0.45   |
| Share of natives                    |                                             | -0.7832       | -6.07   | -0.7681  | -5.95   |                                         | -0.4661       | -2.78   | -0.4654  | -2.77   |
| Share of total employment           |                                             | 1.3186        | 2.34    | 2.0024   | 3.30    |                                         | 0.8158        | 1.50    | 1.0016   | 1.64    |
| <b>Sales occupation</b>             | 0.076                                       |               |         |          |         | 0.085                                   |               |         |          |         |
| Number of established co-ethnics    |                                             | <b>0.0076</b> | 4.72    | 0.0077   | 4.76    |                                         | <b>0.0073</b> | 4.77    | 0.0073   | 4.87    |
| Number of all established employees |                                             |               |         | -0.0001  | -1.79   |                                         |               |         | -0.00002 | -0.45   |
| Share of natives                    |                                             | -0.4957       | -5.39   | -0.4877  | -5.34   |                                         | -0.4917       | -2.93   | -0.4916  | -2.93   |
| Share of total employment           |                                             | 0.8346        | 2.47    | 1.2715   | 3.24    |                                         | 0.8606        | 1.53    | 1.0580   | 1.66    |
| <b>Machine operative occupation</b> | 0.052                                       |               |         |          |         | 0.073                                   |               |         |          |         |
| Number of established co-ethnics    |                                             | <b>0.0054</b> | 5.37    | 0.0055   | 5.48    |                                         | <b>0.0064</b> | 4.41    | 0.0064   | 4.41    |
| Number of all established employees |                                             |               |         | -0.00004 | -1.93   |                                         |               |         | -0.00002 | -0.43   |
| Share of natives                    |                                             | -0.3514       | -6.89   | -0.3464  | -6.74   |                                         | -0.4290       | -2.75   | -0.4290  | -2.75   |
| Share of total employment           |                                             | 0.5916        | 2.03    | 0.9031   | 2.78    |                                         | 0.7509        | 1.56    | 0.9234   | 1.56    |
| <b>Elementary occupation</b>        | 0.292                                       |               |         | 0.292    | 7.83    | 0.085                                   |               |         |          |         |
| Number of established co-ethnics    |                                             | <b>0.0226</b> | 5.46    | 0.0229   | 5.54    |                                         | <b>0.0073</b> | 3.65    | 0.0073   | 3.69    |
| Number of all established employees |                                             |               |         | -0.0002  | -1.87   |                                         |               |         | -0.00002 | -0.45   |
| Share of natives                    |                                             | -1.4647       | -7.19   | -1.4422  | -7.03   |                                         | -0.4918       | -3.07   | -0.4911  | -3.06   |
| Share of total employment           |                                             | 2.4660        | 2.29    | 3.7599   | 3.12    |                                         | 0.8609        | 1.56    | 1.0570   | 1.68    |
| <b>Other controls</b>               |                                             |               |         |          |         |                                         |               |         |          |         |
| Education                           |                                             | Yes           |         | Yes      |         |                                         | Yes           |         | Yes      |         |
| Male                                |                                             | Yes           |         | Yes      |         |                                         | Yes           |         | Yes      |         |
| Age                                 |                                             | Yes           |         | Yes      |         |                                         | Yes           |         | Yes      |         |
| Year                                |                                             | Yes           |         | Yes      |         |                                         | Yes           |         | Yes      |         |
| <b>Observations</b>                 |                                             |               | 4,334   |          |         |                                         |               | 1,062   |          |         |

Estimation includes year dummies, a dummy for London and country group dummies – the estimation would not converge with individual countries of birth dummies so 16 regions of birth dummies were used instead.

### 3.7 Conclusion

This chapter has empirically examined the implications of the occupational composition of a job seeker's network of social contacts for the occupational outcome of the job seeker, looking at immigrants in the UK. To the best of our knowledge, it is the first to do so in the context of the UK. Additionally, we have examined how these implications vary across immigrants with different durations of stay in the country. The chapter has also built on Patel and Vella (2013), who explore the question in the context of the USA, by examining whether networks cause job seekers to choose the most popular occupation amongst fellow co-ethnics. We extended their approach to examine the direct relationship between the size of the job seeker's network of contacts (i.e. fellow co-ethnics) in a given occupation and the likelihood that the job seeker ends up in that occupation. This is found to vary across occupations. An advantage of our data is that we can observe in the data whether employment was obtained through referral or not. The chapter thereby distinguishes itself from the existing literature as we are able to establish that fellow co-ethnics may indeed serve as a source of job information for other immigrants.

Referrals and transmission of job vacancy information are very common in the labour market because they reduce search costs for both job seekers and employers. In the case of immigrants, one may expect new immigrants to rely on fellow co-ethnics to find employment because they would likely be unfamiliar with the host country's labour market. We find a strong relationship between the number of fellow co-ethnics who are already established in their respective occupations and the probability that a newly employed recent immigrant obtained employment through referral (i.e. family and friends) in that occupation. Based on our preferred specification, a one unit increase in the number of fellow co-ethnics in a given occupation, at the regional level, leads to 2.7 per cent increase in the likelihood of finding employment in that occupation through referrals. The relationship is found to be weaker amongst immigrants with longer durations of stay in the UK. Furthermore, a unit increase in the size of the newly employed recent immigrant's network in a given occupation is associated with up to a 10 per cent increase in their likelihood of choosing that occupation. As in the case of the USA, we also find that new immigrants are likely to choose the most

popular occupation amongst fellow co-ethnics who are already established in their respective occupations.

While immigrants tend to find employment through networks of social contacts regardless of duration of stay in the UK, fellow co-ethnics appear to have a reduced influence on the occupational outcome of immigrants as their duration of stay in the UK lengthens. The size of the most popular occupation amongst established co-ethnics appears to have no effect on the likelihood that a job seeker with over 5 years in the UK chooses this most popular occupation. Furthermore, we find that the likelihood of finding employment through family and friends responds less to the size of fellow established co-ethnics in those occupations for immigrants with 2 or more years of stay in the UK, suggesting that immigrants may diversify their networks of social contacts with time spent in the UK. Such evidence would support a notion of “social assimilation” in the host country.

### 3.8 Appendix

**Table 3.11: Probability of transition in employment status**

| <i>Duration with employer</i>    | <i>Not in employment 3 months after</i> |
|----------------------------------|-----------------------------------------|
| less than 3 months               | 0.05                                    |
| 3 months but less than 6 months  | 0.03                                    |
| 6 months but less than 12 months | 0.02                                    |

**Table 3.12: Share of recent employees who found employment by “hearing from someone who worked there” (LFS, 2005 onward)**

|                                    | <i>With employer for 3 months or less</i> | <i>With employer for 3-12 months</i> |
|------------------------------------|-------------------------------------------|--------------------------------------|
| <b>Immigrants with stay of:</b>    |                                           |                                      |
| <i>Less than 2 years in the UK</i> | 0.28                                      | 0.27                                 |
| <i>2 - 5 years in the UK</i>       | 0.25                                      | 0.26                                 |
| <i>6 - 10 years in the UK</i>      | 0.22                                      | 0.23                                 |
| <i>Over 10 years in the UK</i>     | 0.27                                      | 0.30                                 |

**Table 3.13: Likelihood of choosing a given occupation – average marginal effect (AME) of the migrant's network of social contacts from equation (3.6.5)**

|                                             | <i>Immigrants with less than 2 years in the UK</i> | <i>Immigrants with over 10 years in the UK</i> |
|---------------------------------------------|----------------------------------------------------|------------------------------------------------|
| <b>Managerial occupation</b>                | 0.0056<br>[0.0055 0.0057]                          | 0.0082<br>[0.0079 0.0086]                      |
| <b>Professional occupation</b>              | 0.0109<br>[0.0107 0.0112]                          | 0.0108<br>[0.0103 0.0112]                      |
| <b>Associate prof. occupation</b>           | 0.0111<br>[0.0109 0.0113]                          | 0.0091<br>[0.0087 0.0094]                      |
| <b>Clerical and secretarial occupation</b>  | 0.0078<br>[0.0076 0.008]                           | 0.0102<br>[0.0098 0.0106]                      |
| <b>Craft and skilled trade occupation</b>   | 0.0063<br>[0.0061 0.0065]                          | 0.0046<br>[0.0043 0.0049]                      |
| <b>Personal &amp; Protective occupation</b> | 0.0095<br>[0.0093 0.0097]                          | 0.0084<br>[0.008 0.0087]                       |
| <b>Sales occupation</b>                     | 0.0065<br>[0.0063 0.0066]                          | 0.0054<br>[0.0051 0.0056]                      |
| <b>Machine ops. occupation</b>              | 0.0069<br>[0.0067 0.0071]                          | 0.0059<br>[0.0056 0.0062]                      |
| <b>Elementary occupation</b>                | 0.0168<br>[0.0165 0.0171]                          | 0.0105<br>[0.0101 0.0109]                      |
| <b>Observations</b>                         | 4,334                                              | 1,062                                          |

*Estimation includes year dummies, a dummy for London and country group dummies – the estimation would not converge with individual countries of birth dummies so 16 regions of birth dummies were used instead. 95% in brackets*

## **Chapter 4: The impact of misclassification error on the measurement of the labour market performance of migrants**

## 4.1 Introduction

Binary explanatory variables are widely used in socio-economic studies. Because they only take a value of 0 or 1, however, measurement error in these regressors (i.e. misclassification error) violates the classical assumption because the measurement error is not uncorrelated with the true measure, nor is its expected value zero (Aigner, 1973). This chapter examines the implications of misclassification error for parameter estimates of interest, using analyses of the relative labour market performance and educational attainments of first and second generation immigrants to the UK, in comparison to natives, for illustration. In this literature, because the survey commonly used for analyses, i.e. the LFS, does not contain information on parental country of birth, researchers have often classified individuals of non-White ethnicities born in the UK as second generation immigrants although they could potentially be the children or grandchildren of UK-born individuals. On the other hand, individuals of White ethnicity with both parents born outside the UK are often classified as natives because, despite being second generation immigrants by definition, they cannot be distinguished from the natives in the data. Hence these studies run the risk of losing the generational element of the assimilation process, arguably a key question of interest.

We observe from the UK General Household Survey (GHS) that over 10 per cent of the UK-born ethnic minority population have parents who were themselves born in the UK and individuals of White ethnicity with both parents born outside the UK represent about 1.5 per cent of UK-born individuals of White ethnicity. This suggests that the categorisation adopted when using LFS data would very likely lead to individuals being assigned the incorrect nativity status. Such classification errors, even at a small magnitude, can have serious consequences for the identification of the regression coefficients (Kreider, 2010).

The starting points of the present chapter are the findings in Algan, Dustmann, Glitz and Manning (2010) [referred to as ADGM hereafter] using the LFS over the period of 1993 to 2007. ADGM compares the educational attainments, employment probabilities and earnings of first and second generation immigrants to those of natives; separate regressions are run when comparing outcomes of

each generation of immigrants to natives, and differences in these outcomes are captured by a nativity variable, which takes a value of 0 or 1.

Because the GHS contains information on parental country of birth of respondents in addition to similar information as the LFS, we replicate the findings reported in ADGM, estimating coefficients of interest using the GHS as an alternative dataset while proceeding with the same method of categorisation used with LFS data. Next, we use information on parental country of birth of respondents from the GHS to obtain a more accurate nativity status of individuals, re-estimate our coefficients of interest and examine the possibility of generalisation from other authors' findings. Such an approach also allows one to obtain an indication of the levels of the potential bias arising from using an indirect measure (Bound et al, 2001) – this also means that we are able to compare outcomes of second generation immigrants of White ethnicity to those of the natives. Our findings confirm that errors caused by classifying individuals as second generation based on birth in the UK combined with ethnicity leads to serious bias. These are not just attenuation biases arising from the classical measurement error; the bias is in both directions and may over or underestimate coefficients by up to 30 per cent. When using information on parental country of birth to measure nativity status, we define individuals with a UK-born parent as natives, those with both parents born outside the UK as second generation immigrants, and all foreign-born individuals are defined as first generation immigrants.

The sample size from the GHS is smaller in comparison to that from the LFS, which arguably makes the latter preferable to the former. It is also conceivable that the researcher may believe the extent of error of classification when using LFS data would be quite small. In the absence of a validation dataset, one may still be able to gauge the implications of classification errors for parameters of interest by introducing, randomly, small amounts of such errors in the sample (see Kreider (2010)). The LFS data used here allow for potential errors in classifying UK-born respondents from non-White ethnicities as second generation when they may in fact be natives. To gauge the extent of the implications of errors of classifications for inferences using LFS data, we randomly allow 1 per cent of these individuals to be wrongly classified in the LFS data.

Despite this apparently modest misclassification problem however, we provide estimates of coefficients showing similar or higher magnitude of biases obtained earlier using the GHS. Moreover, up to 60 per cent of the coefficient estimates are outside the confidence intervals of estimates obtained when one assumes that all UK-born individuals of non-White ethnicities are second generation immigrants. This is particularly the case for ethnic minority groups with a long history of residence in the UK. Our findings that immigrants and the children born to them in the UK generally face lower earnings and employment probabilities relative to the natives, despite their educational attainments advantage, are in line with findings from other studies. However, some of the generational improvements in labour market outcomes observed amongst some ethnic groups when one uses the LFS and classifies all UK-born of non-White ethnicity as second generation, as is the case in other studies, are no longer clear-cut. Accounting for the possibility of classification errors in the sample, it is possible that second generation immigrants do even better than what one may infer from these findings. However, it is quite likely they have worse labour outcomes, relative to natives, in comparison to the first generation immigrants.

## 4.2 Literature Survey

As noted earlier, the assimilation of immigrants in host country was brought to prominence by Chiswick's 1978 seminal paper on the earnings evolution of immigrants in the USA. Studies on this topic generally point out that immigrants often have inferior labour market outcomes in comparison to natives because they lack host country-specific human capital, which the migrant acquires with time spent in the host country. Similarity of income levels between the source country and the destination country of immigrants also plays an important role in the migrant's labour market outcomes, as discussed in section 2.3 of Chapter 2, and partly explains differences in labour market outcomes across different groups of immigrants. The topic of the assimilation of immigrants and their children in the labour market remains of high interest in the policy agenda in several OECD countries (OECD, 2015).

Bell (1997) studies the evolution of male immigrants' earnings from date of entry into the UK and also examines whether these outcomes differ across cohorts, in

comparison to UK-born individuals. Immigrants are subdivided into three groups: Indians, West Indians, and White Europeans and those from the Old Commonwealth. The paper uses data from the GHS spanning the period of 1973 to 1992. Bell finds that immigrants have more years of education in comparison to UK-born individuals but the magnitudes differ across cohorts, which is attributed to changes in the source countries of immigrants over the years. With regard to measures of assimilation, comparison between UK-born and immigrants is restricted by ethnicity – i.e. only immigrants and UK-born individuals of the same ethnic origins are compared. The author finds that, except for White immigrants from Europe and the Old Commonwealth, most immigrants experience lower wages in comparison to UK-born individuals when they enter the labour market. Similar findings are reported in Clark and Drinkwater (2008) who compare employment/self-employment, unemployment, earnings and hours worked of (first generation) immigrants and the natives. Using data from the LFS, Clark and Drinkwater (2008) focus on the years 2000-3 (pre EU enlargement) and 2004-7 (post EU enlargement). They find that these outcomes vary according to the similarity between their home country's economy and that of the UK. Most groups of immigrants, with the exception of those from Australia and pre-existing member states of the EU, face inferior labour market outcomes in comparison to UK-born individuals. Labour market outcomes are found to be quite heterogeneous amongst recent immigrants and strongly influenced by the country of origin of the immigrants, their English proficiency and where they are educated.

Unlike their parents, the children of immigrants born in the host country are less likely to suffer from lack of host country specific human capital such as language skills, quality of education, etc. Nevertheless, empirical evidence suggests that the heterogeneity in the labour market outcome observed amongst first generation immigrants is reflected amongst second generation immigrants. Using the General Social Survey, and the 1910, 1940 and 1980 censuses data, Borjas (1994) exploit the difference in the source countries of immigrants to show that there exists a correlation between levels of educational attainments and wages observed from one generation of immigrant to the next in the context of the USA. He estimates an intergenerational correlation of parents' reading skills and educational attainments of children at 25 per cent, and that of earnings at 40 to

50 per cent. In the context of Canada, Aydemir et al (2009) estimates intergenerational elasticity in annual earning amongst immigrants in Canada at around 0.18, which they suggest is comparable to that of the general population.

Cross-country comparative studies on the performance of immigrants and the children born to them in the host country include Aydemir and Sweetman (2007), who compare the labour market performance of both first and second generation immigrants in the USA and in Canada. For Europe, ADGM compare labour market outcomes and educational attainment of immigrants and their offspring in France, Germany and the UK, in comparison to their native counterparts. They find that immigrants and the children born to them in the UK generally outperform natives in educational attainment. It is the opposite case in Germany, and the results are not clear-cut for France. The paper further shows that any gap between the first generation of immigrants and the natives narrows down in the second generation. In all countries, most groups of immigrants face lower wages and employment probabilities in comparison to the natives across both generations. There appear to be some intergenerational improvement in the UK, especially amongst females, but not so in France and in Germany.

Because the data used by ADGM for the UK (the LFS), does not contain information on parental country of birth, ADGM use birth in the UK and ethnicity as a proxy measure for the nativity status of the respondent. Using a similar measure, Dustmann and Theodoropoulos (2010) support the findings in ADGM that the children of immigrants in UK do less well in the labour market in comparison to natives. Like these studies, other papers that look at the relative labour outcomes of foreign-born individuals and UK-born individuals of non-White ethnicities (often referred to as second generation) include Heath et al (2008), Heath and Li (2008), and Li (2010).

Platt (2005), however, uses a direct measure of nativity status from the British Household Panel Survey to look at intergenerational mobility amongst certain ethnic groups in the UK. She finds that for men, the outcome of the second generation is linked to both parental occupation/class and ethnicity, and highlights that the occupations of parents post-immigration seems to play an important role on the occupational outcome of the child later in life. For females however, it seems that only the occupation of the parent is strongly linked to the

occupation of the children later in life. Furthermore there appears to be more mobility amongst ethnic minorities in comparison to those of White ethnicity.

Since it will be the basis of comparison in the present essay, it is worth briefly returning to ADGM. For the UK, the paper uses the LFS covering the period of 1993 to 2007 and respondents are subdivided into 7 groups according to their ethnicities: Bangladeshi; Black-African; Black-Caribbean; Chinese; Indian; Pakistani; and White. Those with ethnicity reported as mixed or other Asian / other Black are excluded from the sample<sup>17</sup>. Respondents of White ethnicity born in the UK are classified as natives and all those not born in the UK are classified as first generation immigrants. All UK-born individuals of non-White ethnicities are considered as second generation irrespective of whether or not their parents could possibly have been born in the UK, which would make them of higher generation than second (i.e. natives). When examining the outcomes of interest, the models estimated are 1) a comparison of each foreign-born ethnic group with UK-born of White ethnicity (i.e. those considered as natives), and 2) a comparison of each UK-born of non-White ethnic group (i.e. those considered as second generation immigrants) with UK-born of White ethnicity.

Although the shortcomings of such a measure of nativity have been acknowledged in the literature, the implications for the identification of the parameter(s) of interest have not been fully appreciated. Heath and Li (2008), for example, note “For the most recent period there will also have been third-generation groups, particularly in the case of the Irish and Caribbeans. However, we are not able to distinguish these groups in the LFS, although we can in the GHS, and we have therefore simply amalgamated them with the second generation”.

It is well established in the econometric literature that measurement errors in an explanatory variable biases not just the OLS coefficient of the variable but also that of other explanatory variables. Although it is often the case that the bias is attenuating, under certain conditions the direction of the bias may not be clear, and in the case of categorical variables, “sufficiently high” bias can lead to the

---

<sup>17</sup> The paper is similar to Dustmann and Theodoropoulos (2010), but given its cross-country comparison nature it only focuses on investigating relative gaps across generation, and does not attempt to explain these differences.

wrong sign (Bound et al, 2001). The nature of the measurement errors that would arise here from using ethnicity and birth in the UK as a measure of nativity status differ to the classical measurement error given the variables that capture the immigrant's generation status are dichotomous. As such, the error would be one of classification. Measurement error in binary regressors violates the classical assumption because the measurement error is correlated with the true measure and its expected value is non-zero (Aigner, 1973).

Gunderson and Kreider (2009) illustrate, in the context of the impact of food security on children's health outcomes, how a very small magnitude of classification errors in the sample can increase the level of uncertainty in the estimated coefficients. As little as 2 per cent of such errors in the sample can have dire consequences for the reliability of estimates of the parameter(s) of interest. Kreider (2010) for instance, with data on disability and employment status, shows how only about 1.3 per cent of classification errors in the sample can significantly widen the bounds of the parameter of interest. As notes Kreider (2010) "what may not be fully appreciated, however, is that the extreme nature of the measurement error in a binary regressor can result in severe identification deterioration of the regression coefficients in the presence of very few classification errors. For a binary regressor, measurement error implies that the variable's true value must be the polar opposite of its reported value".

The present paper uses the British General Lifestyle Survey – previously known as the GHS – for a validation check, because in this data the country of birth of respondents' parents is reported. This allows one to gauge the extent of biases that may arise from using the categorisation often adopted in the literature when using LFS data. As stated in the introduction, this essay, when using the GHS, makes a clear distinction between ethnicity and the nativity status of individuals, thereby also providing intergenerational labour market outcomes across all major non mixed-race ethnic groups in the UK

### **4.3 The data**

The LFS was described in more detail in Chapter 2. It is the most comprehensive survey with information on personal characteristics as well as labour market

characteristics of individuals, and has a sample of about 53000 to 60000 private households in UK. However, as mentioned earlier, the LFS does not provide information on the respondent's parental birth country, meaning one has to use an indirect measure in determining who is categorised as second generation immigrant or native. Therefore, to identify second generation immigrants one must rely on ethnicity rather than parents' country of birth. This inevitably leads to classification error. White ethnic group respondents whose parents are born abroad are classified as natives, while non-White ethnic group respondents whose parents are not born abroad are classified as second generation immigrants rather than some higher-order generation.

The General Household Survey (GHS) – now called the General Lifestyle Survey – is also conducted by the ONS and is also intended to be representative of the whole population. The survey was first conducted in 1971, two years before the LFS started. Since then, the survey has been carried out on an annual basis, with a targeted sample of about 15,000 households. In 2007, the survey was renamed the General Lifestyle Survey. The GHS, unlike the LFS, is a much broader survey and thereby includes various socio-demographic variables in comparison to the LFS. Like the LFS, however, the GHS also contains a range of socio-economic variables including employment and education related variables. Addresses contacted for interview were drawn from a stratified sample of postcode sectors. These postcode sectors are the aggregation level before the smallest level of aggregation unit available – i.e. postal addresses, those used for the LFS.

The data from the GHS used in this study covers the period of 1993 to 2006<sup>18</sup>, except for the years 1997 and 1999 for which there are no data available. Having a dataset with parental countries of birth not only permits a more appropriate way of identifying second generation immigrants, but also provides an opportunity to confirm the stability of estimates which are derived from using a proxy measure in the absence of information on parental country of birth as is the case when using the LFS.

Our three outcomes of interest are educational attainments, employment probabilities and earnings. Following ADGM, we use age of completion of continuous full-time education as the measure of educational attainment. As

---

<sup>18</sup> This is the latest year we could obtain data for; data for further years require a licence which was refused.

noted by the authors, age of completion of continuous full-time education (edage) is the best measure of educational attainment in the context of this analysis given that an overwhelming number of foreign qualifications are reported as “other” in the LFS<sup>19</sup>. The GHS reports age of completion of continuous full-time education only for years after 1998, but terminal age of full-time education (tea) is however reported throughout. We use this as our measure of educational attainment whenever age of completion of continuous full-time education is not available. Since this measure may contain intermittent spells of education, the analysis is restricted to those with reported edage/tea values less than or equal to 34<sup>20</sup>.

The main variable used to compute employment status from the LFS is a self-reported economic activity status. It provides the employment status of individuals as well as the reasons for economic inactivity. Those who are reported as being an employee (including unpaid family workers), self-employed or on a government scheme are categorised as employed, and all the remaining individuals are categorised as not employed (no distinction for reasons of economic inactivity is accounted for). The GHS contains similar measure, which we use.

As with many studies, including ADGM, hourly wages are computed by dividing weekly earnings by weekly hours worked. Where an individual has two jobs, information on earnings and hours worked for both jobs are used. This chapter uses the usual hours worked (including overtime) for the week in question with the LFS data. Here, unlike in the case of educational attainments and employment probabilities, information on earnings are only reported in the first and last waves, so the sample will be of a smaller size. In the GHS, however, the only consistent measure of hours worked throughout the dataset is the number of hours worked in the main job. Therefore, analysis using GHS data does not account for second jobs where those exist.

---

<sup>19</sup> This can be at times up to 1/3. Drinkwater et al (2009) highlights similar issues.

<sup>20</sup> It is however worth pointing out that “tea” remains a very good proxy for “edage” when one compares the two measures whenever they are both available. The correlation between the two variables, when both are truncated at age 34, is 100%.

**Table 4.1: Summary statistics of sample, LFS (1993 to 2007)**

| <b>Generation status</b> | <b>Proportion (%)</b> |                 | <b>Age left full-time education</b> |                 |                 |                 | <b>Hourly wages</b> |                 |                 |                 | <b>Employment</b> |                 |                 |                 |
|--------------------------|-----------------------|-----------------|-------------------------------------|-----------------|-----------------|-----------------|---------------------|-----------------|-----------------|-----------------|-------------------|-----------------|-----------------|-----------------|
|                          | <b>Both gender</b>    |                 | <b>Male</b>                         |                 | <b>Female</b>   |                 | <b>Male</b>         |                 | <b>Female</b>   |                 | <b>Male</b>       |                 | <b>Female</b>   |                 |
|                          | 1 <sup>st</sup>       | 2 <sup>nd</sup> | 1 <sup>st</sup>                     | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>     | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>   | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> |
| Natives                  | 90.3                  |                 | 16.9<br>(0.002)                     |                 | 16.9<br>(0.002) |                 | 10.2<br>(0.02)      |                 | 8.1<br>(0.01)   |                 | 78.3              |                 | 66.1            |                 |
| Immigrants               | 8.4                   | 1.8             | 18.6<br>(0.01)                      | 18.6<br>(0.24)  | 18.0<br>(0.01)  | 18.2<br>(0.02)  | 10.9<br>(0.06)      | 9.2<br>(0.10)   | 9.0<br>(0.04)   | 8.6<br>(0.07)   | 73.1              | 59.5            | 55.2            | 53.10           |
| <i>of which</i>          |                       |                 |                                     |                 |                 |                 |                     |                 |                 |                 |                   |                 |                 |                 |
| White                    | 57.5                  | -               | 18.7<br>(0.01)                      |                 | 18.7<br>(0.01)  |                 | 11.9<br>(0.08)      |                 | 9.3<br>(0.05)   |                 | 77.1              |                 | 63.6            |                 |
| Indian                   | 14.9                  | 32.1            | 19.2<br>(0.03)                      | 19.6<br>(0.04)  | 17.8<br>(0.02)  | 18.9<br>(0.04)  | 10.4<br>(0.11)      | 9.6<br>(0.19)   | 8.1<br>(0.08)   | 8.3<br>(0.10)   | 78.8              | 63.1            | 54.9            | 57.68           |
| Pakistani                | 8.5                   | 21.9            | 17.5<br>(0.04)                      | 18.9<br>(0.07)  | 14.7<br>(0.05)  | 17.8<br>(0.06)  | 8.0<br>(0.12)       | 7.9<br>(0.17)   | 7.5<br>(0.18)   | 7.7<br>(0.18)   | 63.8              | 49.8            | 16.4            | 35.11           |
| Black-African            | 7.1                   | 7.6             | 20.8<br>(0.04)                      | 20.2<br>(0.09)  | 18.9<br>(0.04)  | 19.6<br>(0.08)  | 9.3<br>(0.15)       | 9.4<br>(0.25)   | 8.1<br>(0.09)   | 9.0<br>(0.22)   | 61.0              | 60.2            | 47.6            | 55.70           |
| Black-Caribbean          | 5.0                   | 30.8            | 16.5<br>(0.03)                      | 17.4<br>(0.03)  | 16.9<br>(0.02)  | 17.6<br>(0.02)  | 9.1<br>(0.14)       | 9.6<br>(0.20)   | 8.4<br>(0.09)   | 8.9<br>(0.11)   | 64.6              | 65.2            | 60.8            | 61.56           |
| Bangladeshi              | 3.5                   | 4.1             | 16.9<br>(0.06)                      | 18.4<br>(0.190) | 14.6<br>(0.07)  | 17.4<br>(0.22)  | 6.5<br>(0.17)       | 8.1<br>(0.45)   | 7.8<br>(0.32)   | 8.0<br>(0.36)   | 55.8              | 41.3            | 12.8            | 35.18           |
| Chinese                  | 3.5                   | 3.6             | 19.4<br>(0.07)                      | 19.7<br>(0.15)  | 18.7<br>(0.06)  | 19.8<br>(0.13)  | 10.1<br>(0.25)      | 9.0<br>(0.42)   | 9.0<br>(0.20)   | 9.2<br>(0.43)   | 63.1              | 58.6            | 52.1            | 60.59           |

*Note: sample designed to replicate Algan et al (2010). Nativity status obtained using information on ethnicity and birth in the UK. As in the paper, only individuals of age 26+ are kept in the sample for the summary pertaining to educational attainment. Standard error in parenthesis.*

**Table 4.2: Summary statistics of sample, GHS (1993 to 2006)**

| <b>Generation status</b> | <b>Proportion (%)</b> |                 | <b>Age left full-time education</b> |                 |                 |                 | <b>Hourly wages</b> |                 |                 |                 | <b>Employment</b> |                 |                 |                 |
|--------------------------|-----------------------|-----------------|-------------------------------------|-----------------|-----------------|-----------------|---------------------|-----------------|-----------------|-----------------|-------------------|-----------------|-----------------|-----------------|
|                          | <b>Both gender</b>    |                 | <b>Male</b>                         |                 | <b>Female</b>   |                 | <b>Male</b>         |                 | <b>Female</b>   |                 | <b>Male</b>       |                 | <b>Female</b>   |                 |
|                          | 1 <sup>st</sup>       | 2 <sup>nd</sup> | 1 <sup>st</sup>                     | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>     | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>   | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> |
| Natives                  | 89.8                  |                 | 17.3                                |                 | 17.1            |                 | 9.9                 |                 | 7.9             |                 | 80.1              |                 | 67.0            |                 |
|                          |                       |                 | (0.01)                              |                 | (0.01)          |                 | (0.06)              |                 | (0.05)          |                 |                   |                 |                 |                 |
| Immigrants               | 8.4                   | 1.8             | 19.0                                | 19.0            | 18.2            | 18.9            | 10.9                | 9.6             | 8.8             | 8.2             | 76.3              | 64.1            | 57.6            | 57.6            |
|                          |                       |                 | (0.06)                              | (0.15)          | (0.06)          | (0.13)          | (0.21)              | (0.31)          | (0.10)          | (0.17)          |                   |                 |                 |                 |
| <i>of which</i>          |                       |                 |                                     |                 |                 |                 |                     |                 |                 |                 |                   |                 |                 |                 |
| White                    | 54.5                  | -               | 19.0                                |                 | 18.8            |                 | 11.8                |                 | 9.1             |                 | 80.9              |                 | 66.2            |                 |
|                          |                       |                 | (0.09)                              |                 | (0.07)          |                 | (0.24)              |                 | (0.15)          |                 |                   |                 |                 |                 |
| Indian                   | 15.5                  | 31.8            | 19.3                                | 19.4            | 17.7            | 19.5            | 10.5                | 9.6             | 7.7             | 7.6             | 79.9              | 67.3            | 56.2            | 61.7            |
|                          |                       |                 | (0.13)                              | (0.23)          | (0.14)          | (0.24)          | (0.52)              | (0.38)          | (0.17)          | (0.26)          |                   |                 |                 |                 |
| Pakistani                | 9.1                   | 22.1            | 17.7                                | 19.1            | 14.3            | 17.9            | 7.7                 | 8.0             | 7.7             | 7.2             | 65.4              | 55.1            | 18.2            | 35.3            |
|                          |                       |                 | (0.22)                              | (0.34)          | (0.27)          | (0.29)          | (0.31)              | (0.64)          | (0.55)          | (0.42)          |                   |                 |                 |                 |
| Black-African            | 8.4                   | 9.1             | 21.8                                | 21.1            | 19.4            | 21.2            | 9.6                 | 10.8            | 8.6             | 8.6             | 68.4              | 60.7            | 50.6            | 63.3            |
|                          |                       |                 | (0.25)                              | (0.50)          | (0.22)          | (0.44)          | (0.57)              | (1.12)          | (0.29)          | (0.35)          |                   |                 |                 |                 |
| Black-Caribbean          | 5.7                   | 31.2            | 16.8                                | 17.9            | 17.3            | 18.3            | 7.9                 | 9.8             | 8.3             | 9.1             | 66.3              | 67.9            | 66.3            | 67.8            |
|                          |                       |                 | (0.18)                              | (0.22)          | (0.17)          | (0.18)          | (0.32)              | (0.55)          | (0.26)          | (0.31)          |                   |                 |                 |                 |
| Bangladeshi              | 3.1                   | 3.0             | 17.4                                | 18.4            | 15.4            | 18.2            | 6.2                 | 8.2             | 6.2             | 8.4             | 59.4              | 66.7            | 15.7            | 36.8            |
|                          |                       |                 | (0.35)                              | (1.03)          | (0.37)          | (0.83)          | (0.37)              | (1.38)          | (0.60)          | (1.20)          |                   |                 |                 |                 |
| Chinese                  | 3.6                   | 2.8             | 20.5                                | 20.8            | 19.6            | 20.4            | 15.1                | 13.3            | 8.8             | 7.5             | 68.0              | 70.4            | 62.7            | 65.7            |
|                          |                       |                 | (0.46)                              | (0.98)          | (0.39)          | (0.81)          | (3.86)              | (4.72)          | (0.53)          | (0.50)          |                   |                 |                 |                 |

*Note: sample designed to replicate Algan et al (2010). Nativity status obtained using information on ethnicity and birth in the UK. As in the paper, only individuals of age 26+ are kept in the sample for the summary pertaining to educational attainment. Standard error in parenthesis.*

Table 4.3: Breakdown of nativity status using the GHS

| <i><b>Ethnicity</b></i> | <i><b>Both parents born overseas (in %)</b></i> | <i><b>One parent born overseas (in %)</b></i> | <i><b>Both parents born in the UK (in %)</b></i> |
|-------------------------|-------------------------------------------------|-----------------------------------------------|--------------------------------------------------|
| White                   | <b>1.5</b>                                      | 5.4                                           | 93.1                                             |
| Indian                  | 93.0                                            | <b>4.9</b>                                    | <b>2.1</b>                                       |
| Pakistani               | 95.0                                            | <b>4.5</b>                                    | <b>0.5</b>                                       |
| Black-African           | 84.6                                            | <b>9.2</b>                                    | <b>6.2</b>                                       |
| Black-Caribbean         | 78.4                                            | <b>10.3</b>                                   | <b>11.3</b>                                      |
| Bangladeshi             | 95.6                                            | <b>3.3</b>                                    | <b>1.1</b>                                       |
| Chinese                 | 90.5                                            | <b>7.1</b>                                    | <b>2.4</b>                                       |
| <b>Overall share</b>    | <b>3.2</b>                                      | <b>5.5</b>                                    | <b>91.3</b>                                      |

Table 4.4: Source of misclassification

|                                                        | <i><b>White ethnicity</b></i>               |                                       | <i><b>non-White ethnicity</b></i>    |                                             |
|--------------------------------------------------------|---------------------------------------------|---------------------------------------|--------------------------------------|---------------------------------------------|
|                                                        | <i><b>LFS generation status</b></i>         | <i><b>GHS generation status</b></i>   | <i><b>LFS generation status</b></i>  | <i><b>GHS generation status</b></i>         |
| <i><b>Born overseas</b></i>                            | 1 <sup>st</sup> generation immigrant        | 1 <sup>st</sup> generation immigrant  | 1 <sup>st</sup> generation immigrant | 1 <sup>st</sup> generation immigrant        |
| <i><b>Born in the UK, with</b></i>                     |                                             |                                       |                                      |                                             |
| <i><b>Both parents born in UK</b></i>                  | Native (3 <sup>rd</sup> + generation)       | Native (3 <sup>rd</sup> + generation) | <b>2<sup>nd</sup> generation</b>     | <b>Native (3<sup>rd</sup> + generation)</b> |
| <i><b>One parent born abroad</b></i>                   | Native (3 <sup>rd</sup> + generation)       | Native (2.5 <sup>th</sup> generation) | <b>2<sup>nd</sup> generation</b>     | <b>Native (2.5<sup>th</sup> generation)</b> |
| <i><b>Both parents born abroad</b></i>                 | <b>Native (3<sup>rd</sup> + generation)</b> | <b>2<sup>nd</sup> generation</b>      | 2 <sup>nd</sup> generation           | 2 <sup>nd</sup> generation                  |
| <i><b>Proportion of Misclassification: (1.7 %)</b></i> | <b>1.5%</b>                                 |                                       | <b>11.9%</b>                         |                                             |

Sample summaries from the LFS and the GHS are reported in Table 4.1 and Table 4.2. In these tables, nativity status is measured based on the categorisation adopted in the literature when using LFS data. We then discuss the extent of classification errors, which arises from such an approach based on information on parental country of birth from the GHS. The samples in both surveys are restricted to individuals between 16 and 64 years of age as in ADGM. How the sample summary obtained using LFS data in this chapter compares to the one reported in ADGM can be found in the appendix, in Table 4.11.

### 4.3.1 Sample characteristics

The samples designed to replicate ADGM (see Table 4.1 and Table 4.2) consist of around 90 per cent of UK-born individuals of White ethnicity, 8 per cent foreign-born and 2 per cent UK-born from ethnic minorities<sup>21</sup>. Foreign-born of White ethnic origin represent over half of first generation immigrants, and those of Chinese and Bangladeshi ethnic origins are the smallest groups, representing each about 3 per cent. Second generation immigrants are dominated by individuals of Black-Caribbean and Indian ethnic origins, each accounting for approximately 30 per cent of the non-White UK-born population.

The mean school leaving age of UK-born males of White ethnicity is 17 compared to 19 for both first generation immigrants and UK-born individuals of ethnic minority origins. Immigrant groups have at least the same or more years of continuous full-time education completion in both LFS and GHS samples compare to natives, except foreign-born males of Black-Caribbean origin. Similarly to males, with the exception of first generation female immigrants of Bangladeshi and Pakistani ethnic origins, all immigrant groups outperform natives. These patterns are also observed from using GHS data – see Table 4.2. We see from both the GHS and the LFS that there exists a considerable degree of heterogeneity in the mean school leaving age across different groups of immigrants. For instance, looking at males, we note that Black African, Indian and Chinese outperform other immigrants groups as well as the natives, be it amongst first generation immigrants or those born in the UK.

---

<sup>21</sup> The summary presented here using the LFS is similar to the reported summary in Algan et al. Comparison of the two is provided in the appendix (in Table 2.14).

Looking at Table 4.1 and Table 4.2, the heterogeneity in employment rates amongst the different groups is apparent too. Employment rates are generally lower for first generation immigrants and UK-born individuals of non-White ethnic origins relative to natives across both genders. Foreign-born individuals of White ethnicity and first generation male immigrants of Indian ethnicity, however, have employment rates comparable to natives. Employment rates are particularly low amongst females of Bangladeshi and Pakistani origins, whether born in the UK or abroad. The figures obtained using the GHS remain generally comparable to those obtained using the LFS, although in rare occasions, some employment rates obtained using the GHS are 5 percentage points higher than those obtained using the LFS.

As with employment rates, mean hourly wages vary significantly across the different groups of interest. Generally, foreign-born women and women born in the UK but of non-White ethnicity have higher mean hourly wages compare to UK-born women of White ethnic origin. For males, however, this holds only for few groups. Here too, the figures in Table 4.2 are comparable to those in Table 4.1 albeit with small differences.

### 4.3.2 Source of misclassification error

Take  $i$ , an individual of non-White ethnicity who has both parents born in the UK, and another respondent,  $j$  who similarly is of non-White ethnicity but with both parents born outside the UK. While the GHS enables the researcher to classify  $i$  as a native (order of generation higher than second) and  $j$  as a second generation immigrant, using the LFS, the researcher would classify both  $i$  and  $j$  as second generation immigrants because of lack of information on parents' country of birth in the LFS. Equally, take  $k$ , an individual of White ethnicity born in the UK to parents who were both born outside the UK. While with the GHS  $k$  would be categorised as second generation immigrant, they would be categorised as a native with the LFS. Table 4.3 and Table 4.4 illustrate the source of misclassification error when using ethnicity and birth in the UK as a measure for generation status, contrasting categorisations of individuals when one uses the GHS versus the LFS.

Using information on parental country of birth from the GHS, this chapter categorises anyone born abroad as a first generation immigrant and any individual born in the UK to parents who were both born abroad as second generation immigrant. Those with only one foreign-born parent are viewed as 2.5<sup>th</sup> generation and those with both parents born in the UK are considered 3<sup>rd</sup>+ generations; these two groups of individuals are considered as natives in the present chapter<sup>22</sup>. While Table 4.3 provides detailed breakdown of nativity status, the last row of Table 4.4 shows overall proportions of individuals that are likely to be misclassified if one were to adopt the categorisation often used with LFS data. Respondents of non-White ethnicities classified as second generation while they are in fact natives are more than 11 per cent of UK-born individuals of non-White ethnicities. Equally, those of White ethnicity that are classified as natives while in fact being of second generation are about 1.5 per cent of UK-born individuals of White ethnicity.

The break-down of nativity status of individuals born in the UK using GHS data reveals that the ethnic group with the highest share of individuals born to UK-born parents are Black-Caribbean, with over 11 per cent of individuals having both parents born in the UK – and over 10 per cent have one parent born in the UK. For those of Black-African origin, over 9 per cent have at one parent born in the UK and over 6 per cent have both parents born in the UK. Amongst individuals of Chinese origin, about 7 per cent have one UK-born parent and those with both UK-born parents account for over 2 per cent. Similarly, amongst those of Indian origin, over 2 per cent have both parents born in the UK and another 5 per cent with one parent born in the UK. Amongst individuals of Bangladeshi origin, over 3 per cent have one UK-born parent and about 1 per cent of them have both parents born in the UK. And for UK-born individuals of Pakistani origin, only 0.5 per cent of them have both parents born in the UK, but 4.5 per cent have one parent who is UK-born. It is therefore apparent from Table 4.3 that the categorisation adopted when using LFS data would lead to non-trivial amount of classification errors in the sample.

---

<sup>22</sup> This is motivated by the fact that the characteristics of the 2.5<sup>th</sup> generation, who are included in the 2<sup>nd</sup> generation group often, are different to those of the 2<sup>nd</sup> generation, consistent with remarks made in Ramakrishnan (2004) for the U.S. – see table 2.23 in the appendix for detail.

## 4.4 Methodology

### 4.4.1 Implications of misclassification

One of the earliest studies that examines the consequence of misclassification error in regression analysis is Aigner (1973). The paper considers the case where a binary regressor is measured with error. Aigner shows how error of classification in the regressor would lead to inconsistent estimate of interest, with the resulting estimate being asymptotically biased toward 0. Bollinger (1996), an extension of Aigner (1973), perhaps more clearly illustrate the nature of the attenuation.

Consider the following linear models:

$$Y = \alpha + \beta Z + u \quad (4.4.1)$$

$$Y = \partial + \gamma X + e \quad (4.4.2)$$

with  $X = Z + \zeta$ , where  $\zeta$  is error term.

$Z$  and  $X$  are binary outcomes while  $Y$  is a continuous measure. Equation (4.4.1) is the true model of interest to the researcher, but  $Z$  may not be fully observable and is potentially measured with error. The observed measure of  $Z$  is  $X$ . The researcher is interested in estimating the population parameters  $(\alpha, \beta)$  but they are only able to estimate  $(\partial, \gamma)$ .

Letting  $P(X = 1, Z = 0) = p$  and  $P(X = 0, Z = 1) = q$ , and denoting  $P_X$  the marginal probability of  $X=1$ ,  $d$  the inverse of the slope of the least square of  $X$  on  $Y$ , Bollinger (1996) show that without loss of generality, for  $\beta > 0$ , it is bounded by:

$$0 < \gamma \leq \beta \leq \max\{d \cdot P_X + \gamma \cdot (1 - P_X), d \cdot (1 - P_X) + \gamma \cdot P_X\} \quad (4.4.3)$$

It is hence clear from equation (4.4.3) that the true coefficient estimate of  $\beta$  would be underestimated by equation (4.4.2), that is the estimate of  $\gamma$  would be an underestimate of  $\beta$ . Bollinger (1996) further shows that coefficient estimate on the other co-variate will also be biased in the case of a multivariate regression.

Often however, as in our case, the regression may contain more than one dichotomous explanatory variable with measurement error. Examining the direction of the inconsistency becomes a more complex issue (Cameron and Trivedi (2009)). The case of measurement error in two (continuous) regressors is examined in Garber and Klepper (1980). They suggest that “the coefficient of a

mis-measured regressor is not necessarily attenuated, but at least one of them would be. Klepper (1988) examine identification of coefficients of interest when  $k$  explanatory variables are measured with error and where one or more of these variables is dichotomous. The regression considered in Klepper (1988) is as follows:

$$Y = \alpha + \beta X^* + u \quad (4.4.4)$$

where  $X^* \equiv (x_1^*, x_2^*, \dots, x_k^*)'$  is a  $K \times 1$  vector of true regressors not observed by the analyst and  $x_1^*$  is a dichotomous variable. Instead,  $X \equiv (x_1, x_2, \dots, x_k)'$  is observed.  $\beta \equiv (\beta_1, \beta_2, \dots, \beta_k)'$  is a  $K \times 1$  vector of coefficients of interest and  $u$  is a vector of disturbances. Klepper (1988) denotes the vector of coefficients of the regression of  $Y$  on  $X$  as  $b^D$  and the vector of coefficients of the regression of  $x_i$  on  $Y$  and the remaining regressors as  $b^i$ , and two theorems are derived:

(1): Suppose  $x_1, x_2, \dots, x_k$  may be measured with error and  $x_1$  is defined so that  $b_1^D > 0$ . If  $b^D, b^1, b^2, \dots, b^k$  are not contained in the same orthant, then the set of feasible values of  $\beta_i^L = -\infty$  and  $\beta_i^U = +\infty$  for all  $i$ ;

and

(2): If  $b^D$  and  $b^i, i = 1, 2, \dots, K$ , are all contained in the same orthant, then for all  $i = 1, 2, \dots, K$ ,  $\min(b_i^D, b_i^1, \dots, b_i^k) \leq \beta_i \leq \max(b_i^D, b_i^1, \dots, b_i^k)$ .

Note that  $b_i^k$  is the coefficient estimate on the  $i^{th}$  covariate in the regression of  $x_k$  on  $Y$  and the remaining regressors.  $b_i^D$  is the coefficient estimate on the  $x_i$  in the regression of  $Y$  on  $X$ .

Theorems (1) and (2) entail that the bias may no longer be an attenuation bias because the coefficient estimate on the dichotomous variable (and the other variables) measured with error no longer provide a lower bound for the coefficient estimate of the true measure. The implications are shown to be similar in cases where more than one dichotomous regressor are measured with error. The paper notes that these results may also hold in non-linear regression settings.

Kreider (2010) examine the implications of misclassification error in a regression with a binary dependent variable and two binary regressors, where only one of the regressor suffers from measurement error. The error is allowed to be

arbitrary under very relaxed assumptions – i.e. only a limit is set on the level of classification errors in the data. All 8 possible types of observations containing a row with or without a classification error are considered. Kreider shows that the coefficient estimate on the true variable which suffers from measurement error is a function of the first and second moments of all three variables. He then illustrates how misclassification levels of as small as 2 per cent can bias the coefficients of interest. These are not attenuation biases, they can take any direction. Kreider further provides an empirical illustration with a sample containing 1.3 per cent of classification errors. The findings suggest that this can lead to a 10 percentage points of uncertainty in the coefficient estimates of interest, and the resulting estimates suggest a bias ranging from -35 per cent to 34 per cent.

## 4.5 Empirical Analysis

Given measures outlined in the data section are descriptive statistics, we now turn to regression analysis to control for factors such as region of residence, and potential experience, each of which are known to have influences on our outcome variables of interest.

### 4.5.1 Estimation

The modelling strategy adopted is as follows:

$$Y = \alpha + \beta X + \theta_{eth} D_{eth} + \varepsilon \quad (4.5.1)$$

$Y$  represents different outcome variables of interest, which are educational attainment levels, employment probabilities and earnings.  $X$  is a matrix of control characteristics which vary depending on the outcome variable<sup>23</sup>.  $D_{eth}$  is a vector of binary dummies which capture the difference in intercepts between the base group, i.e. natives, and each ethnic group of immigrants.  $\varepsilon$  is the error term. Separate regressions are run with sample consisting of natives and first

---

<sup>23</sup> Characteristics such as marital status, number of dependent children might also affect labour market outcomes too, but these are not controlled for in Algan et al (2010), the benchmark for our comparison.

generation immigrants, or with the sample consisting of natives and second generation immigrants. The analyses look at each gender separately too.

As both the GHS and the LFS are for the same population and the data generating processes are very similar, the GHS may serve as a good validation dataset of the LFS on the topic of assimilation of immigrants in the UK across two generations. We first estimate our coefficients of interest by opting for an identical strategy of classification of individuals across both the LFS and the GHS. We then use the GHS as a validation dataset given we can now use parental country of birth to, arguably more appropriately, distinguish second generation immigrants. As noted in Bound et al (2001), this strategy of re-estimating our model using the appropriate classification of individuals into native, first and second generation immigrants will give an idea of the nature of the statistical bias from the inferential method used with the LFS.

Given that the LFS data used here allow for potential errors in classifying respondents from non-White ethnicities born in the UK as second generation when they are in fact of higher generation (i.e. natives), we randomly allow for 1 per cent of misclassification in the sub-groups of UK-born individuals of non-White ethnicities and run 10,000 regression with such simulated errors. We then compare the estimates we obtain to those resulting from estimation where we assume all respondents of non-White ethnicities born in the UK are second generation immigrants. In each simulation, we randomly draw 1 per cent of the total observations in each non-White ethnic sub-group and re-classify these individuals as natives before re-estimating our coefficients of interest. Although one could allow for a much higher proportion of misclassification, we conservatively chose 1 per cent to empirically illustrate the potential uncertainty that even such small magnitude of misclassification could introduce in our coefficient estimates.

#### **4.5.1.1 Educational Attainment**

To capture the difference in years of schooling between immigrants and the native population, the age the individual left continuous full-time education is regressed on ethnicity and generational dummies; other covariates include a

second order polynomial in years of birth, region dummies and year dummies. As each group of immigrants is assigned a dummy variable, the reported coefficients ought to capture the difference in conditional means of education levels between immigrants and the native population – and indeed amongst different groups of immigrants. Given the relatively smaller size of second generation immigrants, all individuals still in full-time education are kept in the sample: for these individuals, the dependent variable is right censored at current age as does ADGM. The samples consist of individuals between the ages of 16 and 64.

#### **4.5.1.2 Employment**

Conditional means of employment probabilities are computed using a probit estimator with employment status as the dependent variable and generation status dummies for immigrants as explanatory variables. Other covariates in the regression include region and survey year dummies, educational attainment and potential experience (in quartic form). The sample here, however, excludes individuals reported as being still in full-time education as do ADGM. Estimates of the marginal effects obtained on the dummies are used to compare employment probabilities between immigrants and natives.

#### **4.5.1.3 Earnings**

The differences in earnings between natives and first/second generation immigrants of different ethnic origins are obtained by regressing log hourly wages against ethnicity and generation status dummies, potential experience (in quartic form), region and time dummies. In one set of regressions, educational attainment is not controlled for. In another set of regressions, we control for educational attainment. Doing so allows one to gauge the extent of earnings penalty given immigrants appear to have more years of full-time education in comparison to natives. Hourly wages are adjusted using earning indexes and with 2007 as base year for the summary statistics reported in Table 4.1 and Table 4.2. Separate regressions are run when comparing each generation of immigrants with the reference group.

The direct comparison of the estimates from the replication of ADGM using the LFS and how they compare to the estimates reported in their paper are presented

in the appendix, in Table 4.12 to Table 4.15. One can see that the two are comparable despite slight differences in sample sizes. Thus, the same methodology is now used with GHS data before using information on parental country of birth to measure nativity status.

## 4.6 Results

### 4.6.1 Education

Table 4.5 presents results obtained from the replication of ADGM using LFS data in this chapter. Coefficients obtained from defining second generation as individuals born in the UK to parents who are both foreign-born are presented and contrasted with coefficient estimates obtained from following the categorisation used with LFS data, based on GHS data. The upper half of table contains estimates when the samples consist only of males and the lower half of the table contains estimates when the samples consist only of females.

#### **Males:**

The coefficients obtained in this present paper using the categorisation used in ADGM with both LFS and the GHS show that most foreign-born individuals and UK-born individuals of non-White ethnicities leave full-time education at an older age in comparison to UK-born individuals of White ethnicity – see first 4 columns of results in Table 4.5. Results from the LFS show that first generation immigrants of Black-Caribbean origin and those of Bangladeshi origin, on the other hand, have an average school leaving age of about half a year less than the reference group. Also, individuals of Black-Caribbean origin born in the UK lag behind by about .4 years on average. Results from the GHS point in similar direction, although with coefficients of higher magnitudes at times. The highest educational attainment advantage occurs amongst first generation immigrants of Chinese (Black-African) origin using the LFS (GHS), with a school leaving age of about 3 years more than the reference group.

Using parental country of birth to obtain generation status (i.e. with the GHS) shows that first and second generation immigrants leave full-time education at an older age than natives, except those of Black-Caribbean origin where the second

generation lags behind in educational attainment. However, for all groups with an educational attainment advantage across both generations, we notice a generational deterioration. While they still maintain an educational attainment advantage over the reference group, second generation immigrants tend to leave full-time education at a relatively younger age in comparison to their first generation counterparts.

We observe some noticeable differences for second generation coefficients obtained from the two different methods of categorisation used with GHS data. To illustrate the uncertainty that may result from classification error in the data, we see from Table 4.5 that the coefficient estimate on second generation Black-African is 3.00 – with a 90% confidence interval of [2.49 3.51] – using our proxy measure compared with 3.54 – with a 90% confidence interval of [2.96 4.12] – using parental country of birth, a downward bias of 15 per cent. Other noticeable examples include the change in coefficients on second generation Chinese and second generation Black-Caribbean.

## **Females**

The trend of the relative educational attainment for females is different to that of males. Estimates of interest using LFS data show that first generation immigrants of Pakistani, Black-Caribbean and Bangladeshi origins have less years of schooling in comparison to UK-born individuals of White ethnicity. Amongst UK-born individuals, however, only those of Black-Caribbean and Bangladeshi origin lag behind their native counterparts. Turning to the GHS and adopting the same categorisation as we did with the LFS, both foreign-born and UK-born individuals of Black-Caribbean origin now have an educational attainment advantage over UK-born individuals of White ethnicity, although this is small in magnitude and insignificant for the first generation.

Defining second generation and natives using parental country of birth leads to no change in the direction of the coefficients, but a small change in magnitude is apparent amongst most second generation groups – they become higher. Unlike males however, there is a clear generational improvement across all ethnic groups, except immigrants of White ethnicity, in educational attainment relative to natives. Contrary to the implicit assumption that second generation

immigrants of White ethnicity are similar to natives, we see that they have about half a year of educational attainment advantage in comparison to the natives.

The estimates suggest potential biases, arising from the use of proxy measure, of over 42 per cent downward. The coefficient estimates for second generation of Black-Caribbean origin from using our proxy is 0.29 compared with 0.50 from using parental country of birth. The former has a 90% confidence interval of [0.07 0.5] versus [0.25 0.74] for the latter. In sum, across both genders we see that coefficients obtained using ethnicity and birth in the UK as a proxy measure for nativity status lead to potential biases in estimates.

#### **4.6.2 Employment**

Coefficient estimates capturing differences in employment probabilities between natives and immigrants are reported in Table 4.6. The first 2 columns of results contain estimates obtained from using LFS data. Estimates obtained from using GHS data are presented in next 4 columns, the last 2 of which contain estimates obtained using parental country of birth to categorise individuals as natives, first generation immigrants or second generation immigrants.

##### **Males:**

First and second generation immigrants generally face lower employment probabilities in comparison to natives, regardless of ethnicity. The least disadvantaged group amongst foreign-born is those of White ethnicity, with 4 percentage points lower employment probability than the UK-born of White ethnicity. The most disadvantaged group is those of Black-African ethnicity, with about 20 percentage point lower employment probability in comparison to the UK-born of White ethnicity. Amongst UK-born individuals, the least disadvantage is those of Chinese origin, with a gap of 3 percentage points. The biggest difference occurs amongst those of Pakistani origin, with a 16 percentage point difference. Nevertheless, there is a generational improvement for all groups but Black-Caribbean, with the gaps between UK-born ethnic minorities and UK-born individuals of White ethnicity reducing in magnitude.

When we use parental country of birth to obtain generation status, we see here too a change in coefficient estimates for some second generations immigrants. For example, the coefficient estimate for second generation immigrants of Indian origin is around -0.09 – with a 90% confidence interval of [-0.14 -0.05] – using our proxy measure compared with approximately -0.07 – with a 90% confidence interval of [-0.11 -0.03] – using parental country of birth, around 30 per cent difference in magnitude.

### **Females:**

Amongst females, foreign-born individuals also have lower employment probabilities in comparison to UK-born individuals of White ethnicity. The employment probabilities of those of Pakistani and those of Bangladeshi origins are particularly low. They are respectively 54 and 56 percentage points lower in comparison to the base group. These gaps, albeit still noticeable, are now much lower when comparing UK-born of White ethnicity to those of non-White ethnicity. The biggest improvements are amongst those of Pakistani and those of Bangladeshi origins.

Coefficient estimates obtained from using GHS data paint similar patterns as those obtained from using LFS data. Using parental country of birth to more appropriately re-classify individuals leads to the same conclusion: immigrants and the children born to them in the UK have lower employment probabilities in comparison to individuals born in the UK to both UK-born parents. Amongst the first generation of immigrants, the least disadvantage group is the Black-Caribbean, and the relative gap is particularly pronounced amongst first generation immigrants of Pakistani and Bangladeshi origins, standing at around 50 per cent for the former and around 40 per cent for the latter. These gaps considerably decrease in the second generation although they still remain fairly high – they now stand at 30 per cent and 24 per cent respectively. Another considerable generational improvement occurs amongst the Black-African ethnic groups where the relative gap almost halves across generations.

The implication of classification errors for inference is illustrated by the sharp drop in the estimate for Black-Caribbean, which can be seen if one compares estimates when second generation immigrants are identified using a proxy

measure to those when they are identified using information on parental country of birth. The first estimate is -.04 (with 90% CI of [-.09, -.003] and the second estimate is -.001 (with 90% CI of [-.042, .041]).

### 4.6.3 Earnings

Estimates of coefficients of interest from the regression where educational attainment is controlled for are reported in Table 4.7 and those from the regression where educational attainment is not controlled for are reported in Table 4.8. As with educational attainment, the first 2 columns of results contain estimates from using LFS data, followed by estimates from using GHS data in the next 4 columns.

#### Males:

We see from Table 4.7 that all foreign-born and UK-born individuals of non-White ethnicities have lower hourly wages than UK-born individuals of White ethnicity on average. Looking at the 1<sup>st</sup> column, we see that amongst first generation immigrants, the least disadvantaged are those of White-ethnicity – with a gap slightly below 4 per cent, while all other groups lag behind in earnings by much larger gaps – at least 19 per cent. These negative gaps reduce significantly when one compares UK-born individuals of White ethnicity to those of non-White ethnicities; the earnings growth gap is now at around 24 per cent compared with 56 per cent previously. Turning to coefficient estimates obtained from using GHS data, we see that the coefficients for those of Chinese and Indian ethnic origin are now positive albeit insignificant. UK-born individuals therefore experience less earning disadvantage in comparison to their foreign-born counterparts of the same ethnic origin.

Using parental country of birth to obtain generation status of respondents shows intergenerational improvements in earnings across all ethnic groups. Here, the potential bias is between 16 per cent upward (coefficient for Pakistani) and about 30 per cent downward (coefficient for Black-African).

When educational attainment is not controlled for, we see a decrease in the earnings gaps obtained earlier. This is perhaps expected given that immigrants

generally have higher levels of educational attainment in comparison to the natives. In fact, first generation immigrants of White ethnicity and second generation immigrants of Indian and Chinese ethnic origins experience earning advantages, albeit the estimate is statistically insignificantly different from zero for the latter. All groups show an intergenerational improvement, except for those of Whites ethnic origin where there is a rather very small deterioration (but still positive gap) – as indicated by the last two columns of Table 4.8.

Contrary to the implicit assumption that UK-born individuals of White ethnicity born to immigrant parents have the same labour market outcomes as the natives, the estimates obtained when these individuals are appropriately identified in the GHS indicates that their earnings are different (as in the case for their educational attainments). The implications of classification errors for inference are noticeable here too.

#### **Females:**

Similarly to males, we see from the lower half of Table 4.7 that in comparison to UK-born individuals of White ethnicity, all foreign-born and UK-born individuals of non-White ethnicity have on average lower earnings when one controls for education levels. The relative gap is however smaller for those born in the UK compared to foreign-born individuals. Amongst UK-born, individuals of Black-African ethnic origin have the highest relative earning disadvantage (16 per cent using the LFS and 13 per cent using the GHS). Some of the lowest earnings disadvantage are amongst UK-born females of Pakistani and Bangladeshi origins, albeit statistically insignificant for the latter. Compared with foreign-born of the same ethnicity, they are some the largest improvements. When one measures nativity status based on parental country of birth, these intergenerational improvement become apparent across other ethnic groups too.

Comparing estimates when second generation immigrants are identified using a proxy measure to those when they are identified using information on parental country of birth, here also, one notices differences. For example, estimate on second generation Black-African is -.13 from the former measure versus -.10 from the latter measure, a potential bias of up to 30 per cent. For Black-Caribbean, the estimate changes sharply from -.02 to .003.

Table 4.8 suggests that when educational attainment is not controlled for, the coefficient estimates for most second generation immigrants are positive, implying better relative earnings, similar to what we observed for males. The pattern across the dataset is however not consistent as some groups that show positive gaps when using the LFS now indicate negative gaps when using the GHS, and vice versa. However, very few coefficients are significant for second generations.

**Table 4.5: Age left continuous full-time education – base group: natives**

| Generation      | LFS                |                    | GHS                |                   |                                 |                   |
|-----------------|--------------------|--------------------|--------------------|-------------------|---------------------------------|-------------------|
|                 | With proxy measure |                    | With proxy measure |                   | Using parental country of birth |                   |
|                 | 1 <sup>st</sup>    | 2 <sup>nd</sup>    | 1 <sup>st</sup>    | 2 <sup>nd</sup>   | 1 <sup>st</sup>                 | 2 <sup>nd</sup>   |
| <b>Males</b>    |                    |                    |                    |                   |                                 |                   |
| White           | 1.36***<br>(0.02)  |                    | 1.40***<br>(0.06)  |                   | 1.41***<br>(0.06)               | 0.43***<br>(0.10) |
| Indian          | 2.05***<br>(0.05)  | 1.99***<br>(0.06)  | 1.83***<br>(0.10)  | 1.69***<br>(0.17) | 1.83***<br>(0.10)               | 1.80***<br>(0.17) |
| Pakistani       | 0.56***<br>(0.06)  | 1.36***<br>(0.07)  | 0.34*<br>(0.14)    | 1.03***<br>(0.21) | 0.34*<br>(0.14)                 | 1.11***<br>(0.21) |
| Black-African   | 2.69***<br>(0.07)  | 1.88***<br>(0.13)  | 3.65***<br>(0.16)  | 3.00***<br>(0.31) | 3.65***<br>(0.16)               | 3.54***<br>(0.35) |
| Black-Caribbean | -0.55***<br>(0.05) | -0.40***<br>(0.05) | -0.66***<br>(0.18) | -0.06<br>(0.17)   | -0.66***<br>(0.18)              | -0.03<br>(0.19)   |
| Bangladeshi     | -0.51***<br>(0.10) | 0.53***<br>(0.17)  | -0.26<br>(0.23)    | 1.11*<br>(0.62)   | -0.25<br>(0.23)                 | 1.08*<br>(0.65)   |
| Chinese         | 2.74***<br>(0.11)  | 2.43***<br>(0.18)  | 3.59***<br>(0.24)  | 3.08***<br>(0.65) | 3.59***<br>(0.24)               | 3.51***<br>(0.70) |
| <b>N</b>        | 2,341,463          | 2,197,310          | 72,061             | 67,528            | 71,192                          | 67,528            |
| <b>Females</b>  |                    |                    |                    |                   |                                 |                   |
| White           | 1.41***<br>(0.02)  |                    | 1.45***<br>(0.05)  |                   | 1.47***<br>(0.05)               | 0.51***<br>(0.09) |
| Indian          | 0.72***<br>(0.04)  | 1.50***<br>(0.05)  | 0.40***<br>(0.09)  | 1.71***<br>(0.15) | 0.42***<br>(0.09)               | 1.88***<br>(0.15) |
| Pakistani       | -2.18***<br>(0.08) | 0.44***<br>(0.07)  | -2.78***<br>(0.13) | 0.14<br>(0.17)    | -2.76***<br>(0.13)              | 0.40**<br>(0.18)  |
| Black-African   | 0.95***<br>(0.07)  | 1.47***<br>(0.12)  | 1.53***<br>(0.13)  | 2.98***<br>(0.28) | 1.55***<br>(0.13)               | 3.24***<br>(0.30) |
| Black-Caribbean | -0.20***<br>(0.05) | -0.13***<br>(0.04) | 0.13<br>(0.15)     | 0.29**<br>(0.13)  | 0.15<br>(0.15)                  | 0.50***<br>(0.15) |
| Bangladeshi     | -2.61***<br>(0.10) | -0.39**<br>(0.18)  | -1.74***<br>(0.21) | 0.62<br>(0.45)    | -1.72***<br>(0.21)              | 0.77*<br>(0.46)   |
| Chinese         | 1.90***<br>(0.09)  | 2.39***<br>(0.18)  | 2.77***<br>(0.20)  | 3.62***<br>(0.54) | 2.79***<br>(0.20)               | 4.18***<br>(0.58) |
| <b>N</b>        | 2,356,030          | 2,193,239          | 79,186             | 73,826            | 78,264                          | 73,826            |

Table reports coefficients obtained by regressing age left continuous full-time education on generation and ethnicity dummies in a censored regression; dependent variable capped at current age for individuals still in full-time education. Other covariates include a second order polynomial in years of birth, region dummies and time dummies. \*\*\* indicates statistical significance at 1% level, \*\* at 5% level and \* indicates 10% level. Baseline: Native White Population. Robust standard errors when using GHS data. Standard errors are clustered by individual id for estimation using LFS data. Samples comprise 16-64 year olds.

**Table 4.6: Employment probabilities – base group: natives**

| Generation      | LFS                 |                     | GHS                |                    |                                 |                    |
|-----------------|---------------------|---------------------|--------------------|--------------------|---------------------------------|--------------------|
|                 | With proxy measure  |                     | With proxy measure |                    | Using parental country of birth |                    |
|                 | 1 <sup>st</sup>     | 2 <sup>nd</sup>     | 1 <sup>st</sup>    | 2 <sup>nd</sup>    | 1 <sup>st</sup>                 | 2 <sup>nd</sup>    |
| <b>Males</b>    |                     |                     |                    |                    |                                 |                    |
| White           | -0.04***<br>(0.001) |                     | -0.02*<br>(0.01)   |                    | -0.02*<br>(0.01)                | -0.01<br>(0.01)    |
| Indian          | -0.07***<br>(0.003) | -0.05***<br>(0.005) | -0.06***<br>(0.01) | -0.09***<br>(0.03) | -0.05***<br>(0.01)              | -0.07***<br>(0.03) |
| Pakistani       | -0.17***<br>(0.004) | -0.16***<br>(0.006) | -0.18***<br>(0.02) | -0.18***<br>(0.04) | -0.18***<br>(0.02)              | -0.18***<br>(0.04) |
| Black-African   | -0.23***<br>(0.005) | -0.15***<br>(0.010) | -0.20***<br>(0.03) | -0.23***<br>(0.06) | -0.20***<br>(0.03)              | -0.26***<br>(0.06) |
| Black-Caribbean | -0.08***<br>(0.005) | -0.15***<br>(0.005) | -0.10***<br>(0.03) | -0.15***<br>(0.03) | -0.10***<br>(0.03)              | -0.13***<br>(0.03) |
| Bangladeshi     | -0.19***<br>(0.006) | -0.13***<br>(0.015) | -0.23***<br>(0.04) | -0.01<br>(0.08)    | -0.23***<br>(0.04)              | -0.02<br>(0.09)    |
| Chinese         | -0.08***<br>(0.007) | -0.03***<br>(0.013) | -0.06<br>(0.04)    |                    | -0.06<br>(0.04)                 |                    |
| <b>N</b>        | 2,213,117           | 2,070,377           | 67,718             | 63,303             | 66,896                          | 63,305             |
| <b>Females</b>  |                     |                     |                    |                    |                                 |                    |
| White           | -0.08***<br>(0.002) |                     | -0.04***<br>(0.01) |                    | -0.04***<br>(0.01)              | -0.003<br>(0.01)   |
| Indian          | -0.18***<br>(0.003) | -0.06***<br>(0.005) | -0.14***<br>(0.02) | -0.05*<br>(0.03)   | -0.14***<br>(0.02)              | -0.05<br>(0.03)    |
| Pakistani       | -0.54***<br>(0.003) | -0.30***<br>(0.006) | -0.48***<br>(0.02) | -0.33***<br>(0.03) | -0.48***<br>(0.02)              | -0.34***<br>(0.03) |
| Black-African   | -0.24***<br>(0.005) | -0.13***<br>(0.011) | -0.24***<br>(0.02) | -0.08<br>(0.05)    | -0.24***<br>(0.02)              | -0.07<br>(0.06)    |
| Black-Caribbean | -0.03***<br>(0.005) | -0.07***<br>(0.004) | 0.01<br>(0.02)     | -0.04*<br>(0.02)   | 0.01<br>(0.02)                  | -0.001<br>(0.03)   |
| Bangladeshi     | -0.56***<br>(0.005) | -0.24***<br>(0.016) | -0.54***<br>(0.03) | -0.43***<br>(0.08) | -0.54***<br>(0.03)              | -0.41***<br>(0.08) |
| Chinese         | -0.21***<br>(0.007) | -0.03<br>(0.016)    | -0.11**<br>(0.04)  | 0.002<br>(0.11)    | -0.11**<br>(0.04)               | -0.04<br>(0.12)    |
| <b>N</b>        | 2,219,335           | 2,057,703           | 74,317             | 69,075             | 73,448                          | 69,075             |

Marginal effects (at mean) of coefficients obtained by regressing employment status on generation and ethnicity dummies in a probit model; dependent variable takes value 1 for those in employment and 0 otherwise. Samples comprise 16-64 year olds, excluding individuals still in full-time education. Other covariates include age left full-time education, a quartic in potential experience, region dummies and time dummies. \*\*\* indicates statistical significance at 1% level, \*\* at 5% level and \* indicates 10% level. Baseline: Native White Population. Robust standard errors when using GHS data. Standard errors clustered by individual id in regression using LFS data.

**Table 4.7: Earnings (controlling for education) – base group: natives**

| Generation      | LFS                |                    | GHS                |                    |                                 |                   |
|-----------------|--------------------|--------------------|--------------------|--------------------|---------------------------------|-------------------|
|                 | With proxy measure |                    | With proxy measure |                    | Using parental country of birth |                   |
|                 | 1 <sup>st</sup>    | 2 <sup>nd</sup>    | 1 <sup>st</sup>    | 2 <sup>nd</sup>    | 1 <sup>st</sup>                 | 2 <sup>nd</sup>   |
| <b>Males</b>    |                    |                    |                    |                    |                                 |                   |
| White           | -0.04***<br>(0.00) |                    | -0.01<br>(0.01)    |                    | -0.01<br>(0.01)                 | 0.03<br>(0.02)    |
| Indian          | -0.25***<br>(0.01) | -0.04***<br>(0.01) | -0.19***<br>(0.02) | 0.02<br>(0.03)     | -0.19***<br>(0.02)              | 0.02<br>(0.03)    |
| Pakistani       | -0.35***<br>(0.01) | -0.11***<br>(0.02) | -0.30***<br>(0.03) | -0.14***<br>(0.05) | -0.30***<br>(0.03)              | -0.12**<br>(0.05) |
| Black-African   | -0.38***<br>(0.01) | -0.24***<br>(0.02) | -0.39***<br>(0.04) | -0.13*<br>(0.08)   | -0.39***<br>(0.04)              | -0.18**<br>(0.08) |
| Black-Caribbean | -0.19***<br>(0.01) | -0.11***<br>(0.01) | -0.26***<br>(0.04) | -0.08**<br>(0.04)  | -0.26***<br>(0.04)              | -0.11**<br>(0.04) |
| Bangladeshi     | -0.56***<br>(0.02) | -0.11**<br>(0.04)  | -0.56***<br>(0.08) | -0.01<br>(0.11)    | -0.56***<br>(0.08)              | -0.06<br>(0.11)   |
| Chinese         | -0.29***<br>(0.03) | -0.06<br>(0.03)    | -0.07<br>(0.08)    | 0.10<br>(0.15)     | -0.07<br>(0.08)                 | 0.01<br>(0.16)    |
| <b>N</b>        | 378,963            | 358,023            | 41,477             | 39,064             | 40,957                          | 39,064            |
| <b>Females</b>  |                    |                    |                    |                    |                                 |                   |
| White           | -0.05***<br>(0.00) |                    | -0.03*<br>(0.01)   |                    | -0.03*<br>(0.01)                | 0.002<br>(0.03)   |
| Indian          | -0.22***<br>(0.01) | -0.05***<br>(0.01) | -0.18***<br>(0.02) | -0.07**<br>(0.03)  | -0.18***<br>(0.02)              | -0.06*<br>(0.03)  |
| Pakistani       | -0.21***<br>(0.02) | -0.03*<br>(0.02)   | -0.12*<br>(0.06)   | -0.04<br>(0.05)    | -0.12*<br>(0.06)                | -0.04<br>(0.05)   |
| Black-African   | -0.28***<br>(0.01) | -0.16***<br>(0.03) | -0.19***<br>(0.03) | -0.13***<br>(0.05) | -0.19***<br>(0.03)              | -0.10**<br>(0.05) |
| Black-Caribbean | -0.07***<br>(0.01) | -0.03***<br>(0.01) | -0.10**<br>(0.03)  | -0.02<br>(0.03)    | -0.10**<br>(0.03)               | 0.003<br>(0.04)   |
| Bangladeshi     | -0.19***<br>(0.03) | -0.02<br>(0.03)    | -0.38*<br>(0.16)   | 0.06<br>(0.16)     | -0.38*<br>(0.16)                | 0.06<br>(0.16)    |
| Chinese         | -0.19***<br>(0.02) | -0.01<br>(0.04)    | -0.15**<br>(0.05)  | 0.03<br>(0.05)     | -0.15**<br>(0.05)               | 0.01<br>(0.06)    |
| <b>N</b>        | 400,028            | 378,153            | 44,222             | 41,733             | 43,677                          | 41,733            |

Table reports coefficients obtained by regressing log hourly wages on generation and ethnicity dummies in an OLS; other covariates include a quartic in potential experience, region dummies and time dummies. \*\*\* indicates 1% level of significance; \*\* indicates 5% level of significance; and \* indicates 10% level of significance. Baseline: Native White Population. Standard errors clustered by individual id in regression using LFS data. Samples comprise 16-64 year olds

**Table 4.8: Earnings (not controlling for education) – base group: natives**

| Generation      | LFS                |                    | GHS                |                   |                                 |                   |
|-----------------|--------------------|--------------------|--------------------|-------------------|---------------------------------|-------------------|
|                 | With proxy measure |                    | With proxy measure |                   | Using parental country of birth |                   |
|                 | 1 <sup>st</sup>    | 2 <sup>nd</sup>    | 1 <sup>st</sup>    | 2 <sup>nd</sup>   | 1 <sup>st</sup>                 | 2 <sup>nd</sup>   |
| <b>Males</b>    |                    |                    |                    |                   |                                 |                   |
| White           | 0.06***<br>(0.01)  |                    | 0.07***<br>(0.02)  |                   | 0.07***<br>(0.02)               | 0.06*<br>(0.02)   |
| Indian          | -0.09***<br>(0.01) | 0.05***<br>(0.01)  | -0.08**<br>(0.03)  | 0.07<br>(0.04)    | -0.08**<br>(0.03)               | 0.08*<br>(0.04)   |
| Pakistani       | -0.27***<br>(0.02) | -0.04**<br>(0.02)  | -0.27***<br>(0.04) | -0.12**<br>(0.05) | -0.26***<br>(0.04)              | -0.10*<br>(0.05)  |
| Black-African   | -0.21***<br>(0.01) | -0.13***<br>(0.03) | -0.20***<br>(0.04) | 0.001<br>(0.07)   | -0.20***<br>(0.04)              | -0.04<br>(0.07)   |
| Black-Caribbean | -0.20***<br>(0.02) | -0.17***<br>(0.01) | -0.27***<br>(0.05) | -0.09**<br>(0.04) | -0.27***<br>(0.05)              | -0.13**<br>(0.04) |
| Bangladeshi     | -0.57***<br>(0.03) | -0.12**<br>(0.05)  | -0.58***<br>(0.09) | 0.04<br>(0.14)    | -0.58***<br>(0.09)              | -0.03<br>(0.14)   |
| Chinese         | -0.09***<br>(0.03) | 0.04<br>(0.03)     | 0.11<br>(0.08)     | 0.16<br>(0.15)    | 0.12<br>(0.08)                  | 0.11<br>(0.17)    |
| <b>N</b>        | 378,963            | 358,023            | 41,477             | 39,064            | 40,957                          | 39,064            |
| <b>Females</b>  |                    |                    |                    |                   |                                 |                   |
| White           | 0.05***<br>(0.00)  |                    | 0.05***<br>(0.01)  |                   | 0.05***<br>(0.01)               | 0.02<br>(0.03)    |
| Indian          | -0.13***<br>(0.01) | 0.03**<br>(0.01)   | -0.12***<br>(0.03) | -0.03<br>(0.04)   | -0.12***<br>(0.03)              | -0.01<br>(0.04)   |
| Pakistani       | -0.15***<br>(0.02) | 0.01<br>(0.02)     | -0.11<br>(0.06)    | -0.06<br>(0.05)   | -0.11<br>(0.06)                 | -0.04<br>(0.05)   |
| Black-African   | -0.17***<br>(0.01) | -0.07***<br>(0.03) | -0.08**<br>(0.03)  | -0.01<br>(0.04)   | -0.08**<br>(0.03)               | 0.02<br>(0.04)    |
| Black-Caribbean | -0.08***<br>(0.01) | -0.05***<br>(0.01) | -0.08*<br>(0.04)   | -0.01<br>(0.03)   | -0.08*<br>(0.04)                | 0.03<br>(0.04)    |
| Bangladeshi     | -0.15***<br>(0.04) | -0.02<br>(0.04)    | -0.39*<br>(0.17)   | 0.12<br>(0.16)    | -0.39*<br>(0.17)                | 0.12<br>(0.16)    |
| Chinese         | -0.02<br>(0.02)    | 0.11**<br>(0.04)   | 0.05<br>(0.05)     | 0.11<br>(0.06)    | 0.05<br>(0.05)                  | 0.12<br>(0.07)    |
| <b>N</b>        | 400,028            | 378,153            | 44,222             | 41,733            | 43,677                          | 41,733            |

Table reports coefficients obtained by regressing log hourly wages on generation and ethnicity dummies in an OLS; other covariates include a quartic in potential experience, region dummies and time dummies. \*\*\* indicates 1% level of significance; \*\* indicates 5% level of significance; and \* indicates 10% level of significance. Baseline: Native White Population. Standard errors clustered by individual id in regression using LFS data. Samples comprise 16-64 year olds

For all the three measures of assimilation of immigrants across two generations with the GHS, we note that using ethnicity and country of birth as a proxy measure for generation status, instead of parental country of birth, leads to biased estimates in cases. The bias can either be downward or upward, consistent with econometric theory, with over 30 per cent in magnitudes at times. The LFS remains an attractive survey due to the sample size it offers, as one survey contains over 53,000 households (in contrast to 15,000 in the GHS), and it is conducted on a quarterly basis (as opposed to annually for the GHS) – which has perhaps motivated most researchers on the topic to prefer this survey in comparison to the GHS. In what follows, we quantify how much our estimates

obtained for employment probabilities and earnings of second generation male would change when we allow for the possibility of misclassifying some individuals of non-White ethnicities as second generation immigrants while they are in fact natives.

#### **4.6.4 Allowing for 1 per cent of misclassification error in the LFS data used**

In this section, we proceed with our standard specification but allow for potential misclassification of 1 per cent in each subgroup of non-White ethnic origin, given that the impact of generation status on the outcomes of interest is measured separately across the 6 ethnic groups of interest. The consequences of classification errors are separately examined in employment probabilities and earnings regressions. We restrict our sample to males between the ages of 21 and 35. The first and foremost reason for choosing this age group is that they are arguably the most likely to be affected by errors of classification of interest here.

We estimate 10,000 regressions in which 1 per cent of individuals from each of the 6 non-White ethnic origins are randomly allowed to have been misclassified. Maxima and minima coefficient estimates resulting from these regressions are reported in Table 4.9 and Table 4.10, for employment and earnings regressions respectively. Figure 4.1 and Figure 4.2 provide graphical illustration of the range of uncertainty that as little as 1 per cent of classification errors may generate in coefficients of interest.

In the case of employment probabilities, the differences between the possible maxima or minima and the point estimates from adopting a similar method of categorisation as ADGM range from over 4 to about 11 percentage points. For example, for the coefficient on the Black-Caribbean ethnic group, the maximum value we get from the 10,000 estimations with simulated classification errors in the data is -0.108 (with 95% CI of [-0.126,-0.090]) while the minimum value is -0.202 (with 95% CI of [-0.228,-0.176]); the point estimate from adopting a similar method to ADGM is -0.129, with confidence interval [-0.141, -0.127]. We see similar outcomes for coefficients of other ethnic groups in Table 4.9. Graphical illustration for all groups is provided in Figure 4.1; the vertical lines represent the range of coefficient estimates with simulated errors in the sample, and the solid

dots on the line represent estimates obtained from assuming zero error of classification. The range of uncertainty is illustrated by the length of each vertical lines in the figure.

Figure 4.2 plots the density of coefficient estimates from samples with simulated errors. Note that the areas between the two vertical lines indicate confidence intervals of the point estimates we obtain when one assumes that all UK-born individuals of non-White ethnicities are second generation immigrants. We see here that in almost all cases, a considerable number of estimates from the estimations with simulated errors are outside these confidence intervals. The smallest shares occur with coefficients for Chinese and Bangladeshi ethnicities while the largest share occurs for Black-Caribbean ethnicity. It would appear from these percentages that it is coefficient estimates for ethnic groups with a potentially large number of UK-born individuals born to parents who were themselves born in the UK that are likely to be the most affected from these potential classification errors.

When we compare the relative employment probabilities of first and second generation 21 to 35 year old males, we see from Figure 4.3 that it soon becomes unclear whether second generation immigrants would still have better relative employment probabilities than first generation immigrants in a number of cases. The red dots in the figure represent coefficient estimates obtained from adopting the categorisation commonly used with LFS data for second generation immigrants, and the green dots represent coefficient estimates for first generation immigrants. These two dots are connected by a dashed line, which capture relative gaps in employment probabilities. The grey lines in the figure denotes the range of possible estimates of the coefficients for second generation immigrants if one were to introduce 1 per cent of classification errors in the LFS data used here. We see from Figure 4.3 that the dashed line and the grey line often overlap, but this is not always the case, suggesting that the relative employment probability gaps between first and second generation immigrants would be either wider or narrower than one might normally deduce from study misclassifying all UK born of non-White ethnicity as second generation immigrants.

Turning to the earnings results, we see from Table 4.10 that the potential biases arising from errors of classification are in both directions too. Similar uncertainty

as in the case of the employment probabilities analysis also emerges. Here, however, they appear to be less severe; both the levels of biases and the levels of uncertainties are generally much smaller. While the level of uncertainty can be up to 7 percentage points, in most cases it is less than 3 percentage points.

**Table 4.9: Relative Employment Probabilities with 1% misclassification error in the LFS**

|                        | Assuming 0 errors of classification | 10000 simulations with 1% error of classification |                  |
|------------------------|-------------------------------------|---------------------------------------------------|------------------|
|                        |                                     | Maximum estimate                                  | Minimum estimate |
| <b>Indian</b>          | -0.050***                           | -0.041***                                         | -0.088***        |
| Upper Ci               | -0.041                              | -0.028                                            | -0.064           |
| Lower Ci               | -0.059                              | -0.053                                            | -0.112           |
| <b>Pakistani</b>       | -0.110***                           | -0.091***                                         | -0.177***        |
| Upper Ci               | -0.097                              | -0.073                                            | -0.147           |
| Lower Ci               | -0.123                              | -0.110                                            | -0.206           |
| <b>Black-African</b>   | -0.176***                           | -0.149***                                         | -0.262***        |
| Upper Ci               | -0.151                              | -0.113                                            | -0.214           |
| Lower Ci               | -0.200                              | -0.185                                            | -0.310           |
| <b>Black-Caribbean</b> | -0.129***                           | -0.108***                                         | -0.202***        |
| Upper Ci               | -0.117                              | -0.090                                            | -0.176           |
| Lower Ci               | -0.141                              | -0.126                                            | -0.228           |
| <b>Bangladeshi</b>     | -0.050*                             | -0.038**                                          | -0.090***        |
| Upper Ci               | -0.021                              | 0.004                                             | -0.008           |
| Lower Ci               | -0.080                              | -0.080                                            | -0.171           |
| <b>Chinese</b>         | -0.042**                            | -0.031**                                          | -0.076***        |
| Upper Ci               | -0.015                              | -0.065                                            | -0.010           |
| Lower Ci               | -0.069                              | -0.065                                            | -0.142           |
| N                      | 646,923                             | 646,923                                           | 646,923          |

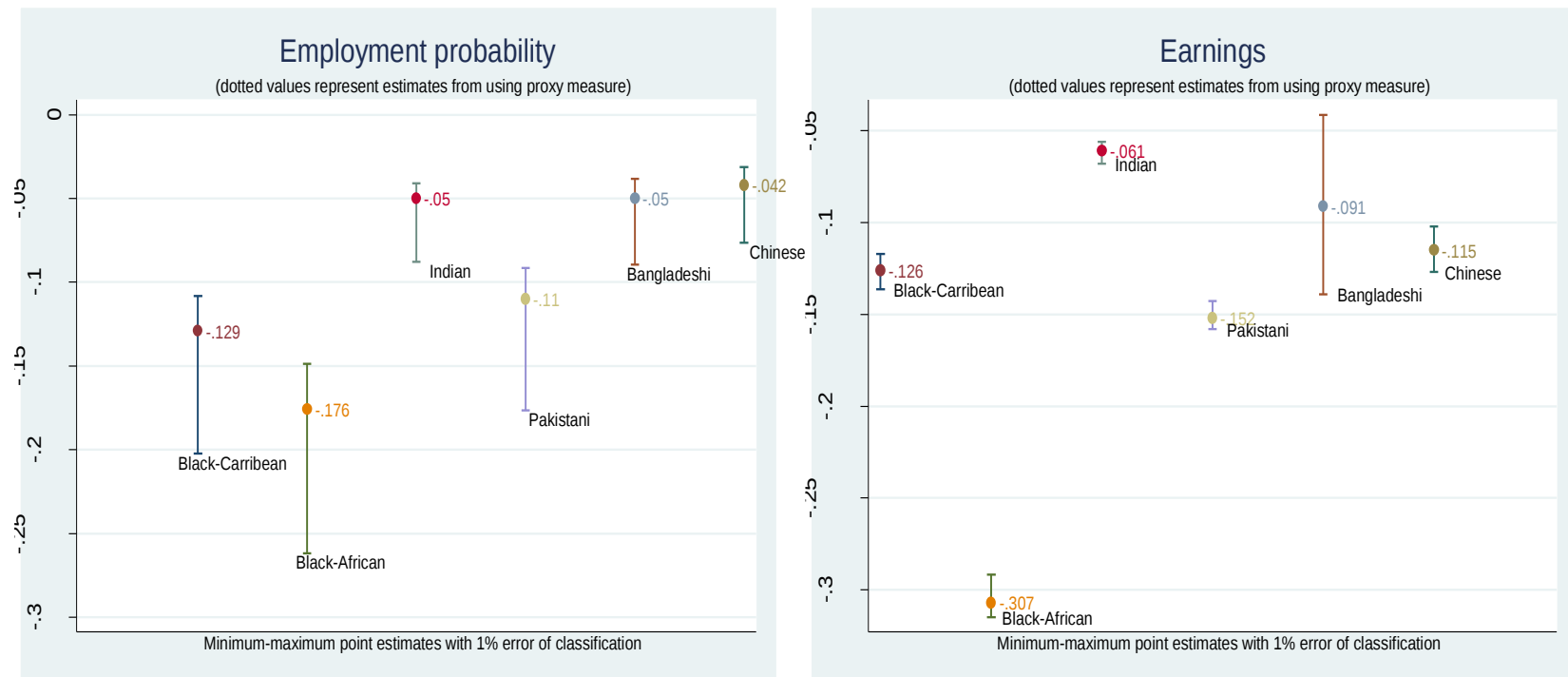
Marginal effects (at mean) of coefficients obtained by regressing employment status on generation and ethnicity dummies in a probit model; dependent variable takes value 1 for those in employment and 0 otherwise. Sample excluding individuals still in full-time education. Other covariates include age left full-time education, a quartic in potential experience, region dummies and time dummies. \*\*\* indicates statistical significance at 1% level, \*\* at 5% level and \* indicates 10% level. Baseline: Native White Population. Standard errors clustered by individual id in regression using LFS data

**Table 4.10: Relative Earnings (controlling for education) with 1% misclassification error in the LFS**

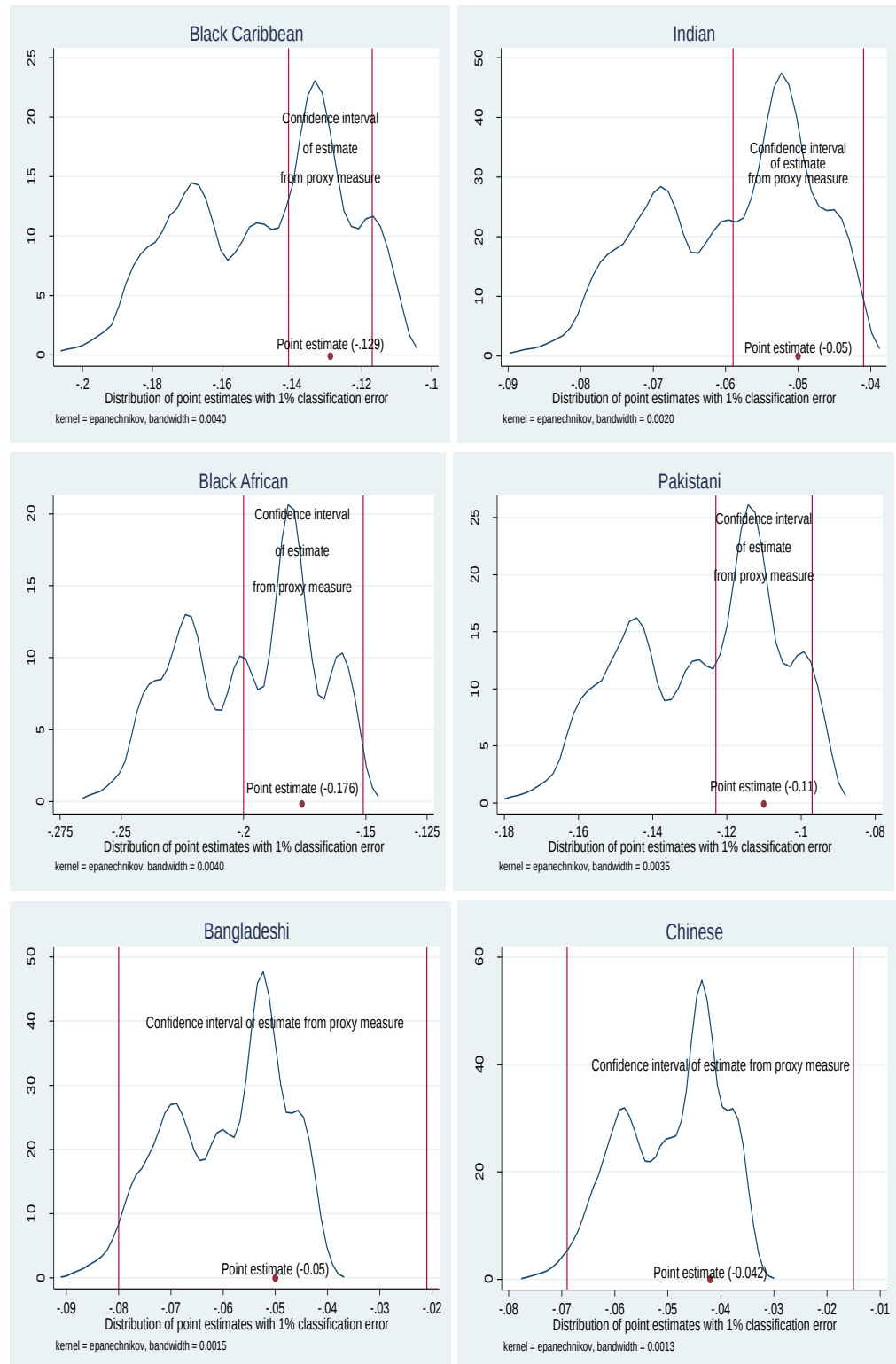
|                        | Assuming 0 errors of classification | 10000 simulations with 1% error of classification |                  |
|------------------------|-------------------------------------|---------------------------------------------------|------------------|
|                        |                                     | Maximum estimate                                  | Minimum estimate |
| <b>Indian</b>          | -0.061***                           | -0.126***                                         | -0.056***        |
| Upper Ci               | -0.033                              | -0.161                                            | -0.028           |
| Lower Ci               | -0.089                              | -0.091                                            | -0.083           |
| <b>Pakistani</b>       | -0.152***                           | -0.158***                                         | -0.143***        |
| Upper Ci               | -0.114                              | -0.120                                            | -0.106           |
| Lower Ci               | -0.190                              | -0.196                                            | -0.180           |
| <b>Black-African</b>   | -0.307***                           | -0.315***                                         | -0.292***        |
| Upper Ci               | -0.246                              | -0.255                                            | -0.234           |
| Lower Ci               | -0.367                              | -0.375                                            | -0.350           |
| <b>Black-Caribbean</b> | -0.126***                           | -0.136***                                         | -0.117***        |
| Upper Ci               | -0.091                              | -0.104                                            | -0.084           |
| Lower Ci               | -0.160                              | -0.169                                            | -0.150           |
| <b>Bangladeshi</b>     | -0.091*                             | -0.139***                                         | -0.041           |
| Upper Ci               | 0.004                               | -0.050                                            | 0.046            |
| Lower Ci               | -0.185                              | -0.228                                            | -0.128           |
| <b>Chinese</b>         | -0.115***                           | -0.127***                                         | -0.102**         |
| Upper Ci               | -0.034                              | -0.048                                            | -0.024           |
| Lower Ci               | -0.197                              | -0.206                                            | -0.180           |
| N                      | 121,003                             | 121,003                                           | 121,003          |

Table reports coefficients obtained by regressing log hourly wages on generation and ethnicity dummies in an OLS, a quartic in potential experience, region dummies and time dummies. \*\*\* indicates 1% level of significance; \*\* indicates 5% level of significance; and \* indicates 10% level of significance. Baseline: Native White Population. Standard errors are clustered by individual id.

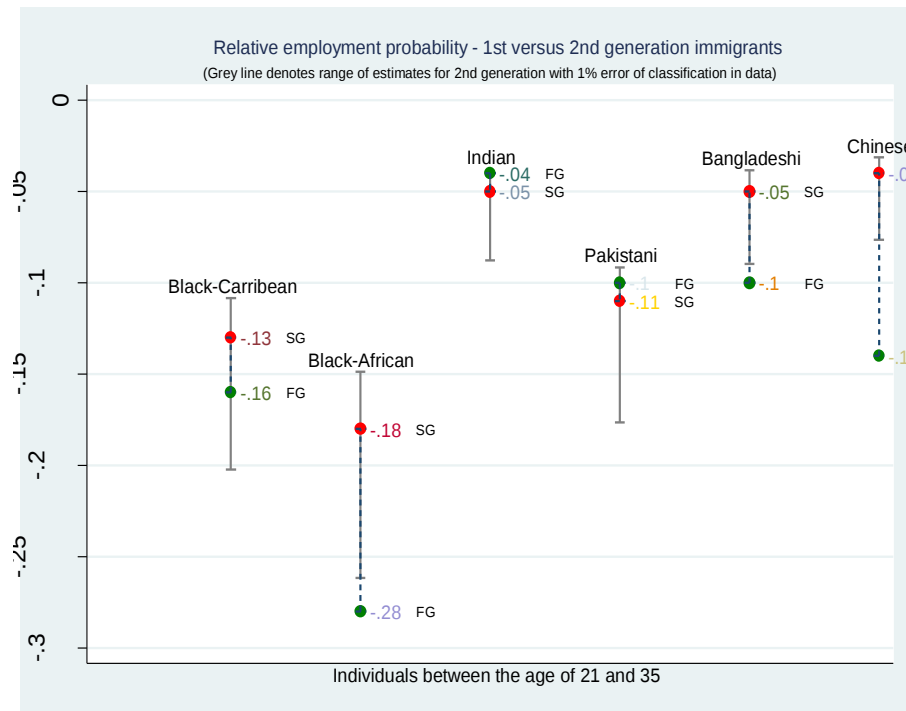
**Figure 4.1: Range of uncertainty arising from allowing 1% misclassification error introduced in the LFS sample**



**Figure 4.2: Relative employment probabilities – densities of point estimates with 1% misclassification error introduced in the LFS sample**



**Figure 4.3: Range of uncertainty when comparing relative employment probabilities of 1<sup>st</sup> and 2<sup>nd</sup> generation immigrants, allowing for 1% misclassification error in LFS data used**



## 4.7 Conclusion

Individuals whose parents were themselves born in the UK account for over 10 per cent of the UK-born ethnic minority population. Similarly, individuals of White ethnicity with both parents born outside the UK represent about 1.5 per cent of UK-born Whites. Because of data limitations, the former group is often mis-classified as second generation immigrants and the latter group is often mis-classified as native in studies using LFS data. These classification errors, even at a small magnitude, can have serious consequences for the identification of the regression coefficients (Kreider, 2010).

In this chapter, we have used an alternative dataset, i.e. the GHS, which contains information on the respondent's parental country of birth, allowing us to obtain labour market outcomes for second generation immigrants across all major ethnic groups in the UK. The outcomes examined are educational attainments, employment probabilities and earnings. The advantage of the GHS is that we are also able to use the categorisation adopted when using LFS data and compare the resulting estimates to those we obtain when using a more accurate measure of who is a second generation immigrant. We find that using birth in the UK and ethnicity to measure who is a second generation immigrant leads to potential biases resulting from the classification errors. Consistent with econometric theory, the direction of the bias is unclear. We find that the biases range from over 30 per cent downward to over 30 per cent upward. It is worth noting that we do not make a distinction of ethnicity amongst the natives although one may argue that labour market outcomes differ across natives of different ethnicities. Comparing labour market outcomes of first and second immigrants to those of natives by ethnic group would address this. However, this is beyond the scope of the current chapter and the GHS data at hand does not contain sufficient observations for each ethnic group to conduct such analysis.

Although the sample size of non-White individuals born in the UK is relatively small in the GHS, these levels of potential bias provide tentative indication of what we can expect for not making a clear distinction between ethnicity and generation status. Allowing for as little as 1 per cent of error in classification of respondents' nativity status amongst non-White ethnic groups in the LFS and running 10,000 regressions with these simulated errors corroborate the findings

from the GHS. Moreover, it further shows a considerable increase in the range of uncertainty of our coefficients. Up to 60 per cent of estimates of interest are outside the confidence intervals of estimates obtained when one assumes all UK-born individuals of non-White ethnicities are second generation immigrants. We conclude that although one may infer from findings using LFS data that there generally is a generational improvement in the labour market outcome of immigrants, the presence of classification errors in the data could lead to a not clear-cut conclusion, especially for ethnic groups with longer history of immigration in the UK: while it is possible that the children of immigrants do even better than what one may infer from findings using LFS data, it also quite likely they have worse labour outcomes, relative to natives, in comparison to first generation immigrants.

## 4.8 Appendix

**Table 4.11: Comparing sample summary to those reported in ADGM**

| Sample Proportion (%)                        |                 |                 | Age Left Full-Time Education |                 |                 |                 | Hourly Wages    |                 |                 |                 | Employment Rate (%) |                 |                 |                 |
|----------------------------------------------|-----------------|-----------------|------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|---------------------|-----------------|-----------------|-----------------|
| LFS: ADGM (2010) – as reported in paper      |                 |                 |                              |                 |                 |                 |                 |                 |                 |                 |                     |                 |                 |                 |
| Generation                                   |                 |                 | Male                         |                 | Female          |                 | Male            |                 | Female          |                 | Male                |                 | Female          |                 |
|                                              | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>              | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>     | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> |
| Natives                                      |                 | 90.3            |                              | 16.9            |                 | 16.9            |                 | 11.1            |                 | 8.5             |                     | 79.0            |                 | 66.5            |
| Immigrants, of which                         | 8.1             | 1.6             | 18.6                         | 18.6            | 17.9            | 18.2            | 11.5            | 10.2            | 9.4             | 9.5             | 73.1                | 59.9            | 55.2            | 53.4            |
| White                                        | 56.8            | -               | 18.6                         |                 | 18.6            |                 | 12.6            | -               | 9.8             |                 | 77.2                | -               | 63.7            | -               |
| Indian                                       | 15.4            | 32.5            | 19.1                         | 19.5            | 17.7            | 18.8            | 11.1            | 10.8            | 8.6             | 9.5             | 78.1                | 63.2            | 55.2            | 57.7            |
| Pakistani                                    | 8.6             | 21.5            | 17.4                         | 18.9            | 14.6            | 17.7            | 8.2             | 8.7             | 7.8             | 8.1             | 63.5                | 49.8            | 16.6            | 35.4            |
| Black African                                | 6.9             | 7.4             | 20.8                         | 20.2            | 18.8            | 19.7            | 9.9             | 10.2            | 8.6             | 10.3            | 60.9                | 60.6            | 47.9            | 56.2            |
| Black Caribbean                              | 5.3             | 31.3            | 16.4                         | 17.4            | 16.8            | 17.6            | 9.9             | 10.3            | 9.2             | 10.0            | 64.7                | 65.9            | 61.1            | 62.2            |
| Bangladeshi                                  | 3.6             | 3.7             | 16.8                         | 18.4            | 14.4            | 17.5            | 6.3             | 8.6             | 7.9             | 8.5             | 55.7                | 43.8            | 12.8            | 36.7            |
| Chinese                                      | 3.5             | 3.7             | 19.4                         | 19.6            | 18.7            | 19.7            | 11.0            | 10.3            | 9.5             | 10.1            | 63.5                | 58.4            | 51.7            | 60              |
| LFS: ADGM (2010) – replication in this paper |                 |                 |                              |                 |                 |                 |                 |                 |                 |                 |                     |                 |                 |                 |
| Generation                                   |                 |                 | Male                         |                 | Female          |                 | Male            |                 | Female          |                 | Male                |                 | Female          |                 |
|                                              | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>              | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> | 1 <sup>st</sup>     | 2 <sup>nd</sup> | 1 <sup>st</sup> | 2 <sup>nd</sup> |
| Native                                       |                 | 90.3            |                              | 16.9            |                 | 16.9            |                 | 10.2            |                 | 8.1             |                     | 78.3            |                 | 66.1            |
| Immigrants, of which                         | 8.1             | 1.6             | 18.6                         | 18.6            | 18.0            | 18.2            | 10.9            | 9.2             | 9.0             | 8.6             | 73.1                | 59.5            | 55.2            | 53.1            |
| White                                        | 57.5            | -               | 18.7                         | -               | 18.7            | -               | 11.9            | -               | 9.3             | -               | 77.1                | -               | 63.6            | -               |
| Indian                                       | 14.9            | 32.1            | 19.2                         | 19.6            | 17.8            | 18.9            | 10.4            | 9.6             | 8.1             | 8.3             | 78.8                | 63.1            | 54.9            | 57.7            |
| Pakistani                                    | 8.5             | 21.9            | 17.5                         | 18.9            | 14.7            | 17.8            | 8.0             | 7.9             | 7.5             | 7.7             | 63.8                | 49.8            | 16.4            | 35.1            |
| Black-African                                | 7.1             | 7.6             | 20.8                         | 20.2            | 18.9            | 19.6            | 9.3             | 9.4             | 8.1             | 9.0             | 61.0                | 60.2            | 47.6            | 55.7            |
| Black-Caribbean                              | 5.0             | 30.8            | 16.5                         | 17.4            | 16.9            | 17.6            | 9.1             | 9.6             | 8.4             | 8.9             | 64.6                | 65.2            | 60.8            | 61.6            |
| Bangladeshi                                  | 3.5             | 4.1             | 16.9                         | 18.4            | 14.6            | 17.4            | 6.5             | 8.1             | 7.8             | 8.0             | 55.8                | 41.3            | 12.8            | 35.2            |
| Chinese                                      | 3.5             | 3.6             | 19.4                         | 19.7            | 18.7            | 19.8            | 10.1            | 9.0             | 9.0             | 9.2             | 63.1                | 58.6            | 52.1            | 60.6            |

*Nativity status obtained using information on ethnicity and birth in the UK. As in the paper, only individuals of age 26+ are kept in the sample for the summary pertaining to educational attainment.*

**Table 4.12: Comparison to ADGM. Age left full-time education; base group: natives, LFS (93 to 07)**

| Generation      | Replication in present paper |                      |                      |                      | ADGM et al (2010)    |                      |                      |                      |
|-----------------|------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                 | Male                         |                      | Female               |                      | Male                 |                      | Female               |                      |
|                 | 1 <sup>st</sup>              | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      |
| White           | 1.358***<br>(0.012)          |                      | 1.408***<br>(0.010)  |                      | 1.335***<br>(0.012)  | -                    | 1.396***<br>(0.010)  | -                    |
| Indian          | 2.053***<br>(0.024)          | 1.985***<br>(0.032)  | 0.717***<br>(0.023)  | 1.503***<br>(0.029)  | 2.017***<br>(0.024)  | 1.958***<br>(0.032)  | 0.692***<br>(0.023)  | 1.494***<br>(0.029)  |
| Pakistani       | 0.560***<br>(0.033)          | 1.361***<br>(0.040)  | -2.180***<br>(0.039) | 0.437***<br>(0.036)  | 0.535***<br>(0.034)  | 1.370***<br>(0.040)  | -2.203***<br>(0.040) | 0.403***<br>(0.036)  |
| Black African   | 2.693***<br>(0.042)          | 1.881***<br>(0.073)  | 0.947***<br>(0.037)  | 1.466***<br>(0.065)  | 2.586***<br>(0.041)  | 1.865***<br>(0.073)  | 0.929***<br>(0.037)  | 1.469***<br>(0.065)  |
| Black Caribbean | -0.550***<br>(0.028)         | -0.397***<br>(0.026) | -0.201***<br>(0.024) | -0.132***<br>(0.023) | -0.547***<br>(0.028) | -0.395***<br>(0.026) | -0.218***<br>(0.023) | -0.128***<br>(0.023) |
| Bangladeshi     | -0.512***<br>(0.054)         | 0.527***<br>(0.092)  | -2.615***<br>(0.053) | -0.392***<br>(0.093) | -0.540***<br>(0.054) | 0.493***<br>(0.094)  | -2.645***<br>(0.053) | -0.330***<br>(0.095) |
| Chinese         | 2.744***<br>(0.061)          | 2.433***<br>(0.095)  | 1.900***<br>(0.049)  | 2.392***<br>(0.094)  | 2.607***<br>(0.060)  | 2.483***<br>(0.099)  | 1.875***<br>(0.049)  | 2.447***<br>(0.096)  |
| N               | 2,341,463                    | 2,197,310            | 2,356,030            | 2,193,239            | 2,226,343            | 2,089,974            | 2,239,887            | 2,085,364            |

Table reports coefficients obtained by regressing age left continuous full-time education on generation and ethnicity dummies in a censored regression; dependent variable capped at current age for individuals still in full-time education. Other covariates include a second order polynomial in years of birth, region dummies and time dummies. \*\*\* indicates statistical significance at 1% level, \*\* at 5% level and \* indicates 10% level. Baseline: Native White Population. Robust standard errors as in ADGM. For the results in the text, standard errors are clustered by individual id.

**Table 4.13: Comparison to ADGM. Employment probability; base group: natives, LFS (93 to 07)**

| Generation      | Replication in present paper |                      |                      |                      | ADGM et al (2010)    |                      |                      |                      |
|-----------------|------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                 | Male                         |                      | Female               |                      | Male                 |                      | Female               |                      |
|                 | 1 <sup>st</sup>              | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      |
| White           | -0.035***<br>(0.001)         |                      | -0.083***<br>(0.002) |                      | -0.039***<br>(0.002) |                      | -0.083***<br>(0.002) |                      |
| Indian          | -0.070***<br>(0.003)         | -0.054***<br>(0.005) | -0.176***<br>(0.003) | -0.058***<br>(0.005) | -0.076***<br>(0.003) | -0.055***<br>(0.005) | -0.177***<br>(0.003) | -0.056***<br>(0.005) |
| Pakistani       | -0.169***<br>(0.004)         | -0.156***<br>(0.006) | -0.540***<br>(0.003) | -0.299***<br>(0.006) | -0.181***<br>(0.004) | -0.166***<br>(0.007) | -0.538***<br>(0.003) | -0.299***<br>(0.007) |
| Black African   | -0.232***<br>(0.005)         | -0.154***<br>(0.010) | -0.243***<br>(0.005) | -0.129***<br>(0.011) | -0.249***<br>(0.005) | -0.159***<br>(0.011) | -0.244***<br>(0.005) | -0.130***<br>(0.011) |
| Black Caribbean | -0.077***<br>(0.005)         | -0.151***<br>(0.005) | -0.032***<br>(0.005) | -0.069***<br>(0.004) | -0.085***<br>(0.005) | -0.158***<br>(0.005) | -0.034***<br>(0.005) | -0.071***<br>(0.005) |
| Bangladeshi     | -0.194***<br>(0.006)         | -0.128***<br>(0.015) | -0.556***<br>(0.005) | -0.243***<br>(0.016) | -0.204***<br>(0.007) | -0.130***<br>(0.016) | -0.559***<br>(0.007) | -0.236***<br>(0.017) |
| Chinese         | -0.082***<br>(0.007)         | -0.033**<br>(0.013)  | -0.210***<br>(0.007) | -0.026<br>(0.016)    | -0.090***<br>(0.007) | -0.035**<br>(0.014)  | -0.214***<br>(0.007) | -0.031*<br>(0.017)   |
| N               | 2,213,117                    | 2,070,377            | 2,219,335            | 2,057,703            | 2,107,340            | 1,971,412            | 2,113,340            | 1,959,190            |

Table reports coefficients obtained by regressing employment status on generation and ethnicity dummies in a probit model; dependent variable takes value 1 for those in employment and 0 otherwise. Sample aged 16 to 64, excluding individuals still in full-time education. Other covariates include age left full-time education, a quartic in potential experience, region dummies and time dummies. \*\*\* indicates statistical significance at 1% level, \*\* at 5% level and \* indicates 10% level. Baseline: Native White Population. Robust standard errors as in ADGM. For the results in the text, standard errors are clustered by individual id.

**Table 4.14: Comparison to ADGM. Log net hourly earnings (controlling for education); base group: natives, LFS (93 to 07)**

| Generation      | Replication in present paper |                      |                      |                      | ADGM et al (2010)    |                      |                      |                      |
|-----------------|------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                 | Male                         |                      | Female               |                      | Male                 |                      | Female               |                      |
|                 | 1 <sup>st</sup>              | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      |
| White           | -0.036***<br>(0.004)         |                      | -0.052***<br>(0.004) |                      | -0.034***<br>(0.004) | -                    | -0.055***<br>(0.004) | -                    |
| Indian          | -0.245***<br>(0.008)         | -0.039***<br>(0.012) | -0.220***<br>(0.008) | -0.047***<br>(0.010) | -0.269***<br>(0.008) | -0.047***<br>(0.012) | -0.236***<br>(0.008) | -0.051***<br>(0.010) |
| Pakistani       | -0.347***<br>(0.013)         | -0.110***<br>(0.015) | -0.209***<br>(0.020) | -0.033**<br>(0.016)  | -0.342***<br>(0.012) | -0.110**<br>(0.016)  | -0.213***<br>(0.020) | -0.039**<br>(0.016)  |
| Black African   | -0.384***<br>(0.013)         | -0.240***<br>(0.023) | -0.283***<br>(0.011) | -0.155***<br>(0.025) | -0.435***<br>(0.014) | -0.301***<br>(0.026) | -0.318***<br>(0.011) | -0.176***<br>(0.026) |
| Black Caribbean | -0.190***<br>(0.014)         | -0.114***<br>(0.012) | -0.074***<br>(0.010) | -0.030***<br>(0.009) | -0.216***<br>(0.014) | -0.128***<br>(0.011) | -0.087***<br>(0.010) | -0.029***<br>(0.009) |
| Bangladeshi     | -0.556***<br>(0.021)         | -0.106**<br>(0.042)  | -0.188***<br>(0.031) | -0.020<br>(0.034)    | -0.553***<br>(0.020) | -0.129***<br>(0.040) | -0.214***<br>(0.034) | -0.038<br>(0.036)    |
| Chinese         | -0.289***<br>(0.024)         | -0.055*<br>(0.033)   | -0.187***<br>(0.019) | -0.011<br>(0.035)    | -0.274***<br>(0.023) | -0.094***<br>(0.038) | -0.173***<br>(0.019) | -0.023<br>(0.034)    |
| N               | 378,963                      | 358,023              | 400,028              | 378,153              | 331,043              | 317,275              | 347,006              | 332,569              |

Table reports coefficients obtained by regressing log hourly wages on generation and ethnicity dummies in an OLS; other covariates include a quartic in potential experience, region dummies and time dummies. \*\*\* indicates 1% level of significance; \*\* indicates 5% level of significance; and \* indicates 10% level of significance. Baseline: Native White Population. Robust standard errors as in ADGM. For the results in the text, standard errors are clustered by individual id.

**Table 4.15: Comparison to ADGM. Log net hourly earnings (not controlling for education); base group: natives, LFS (93 to 07)**

| Generation      | Replication in present paper |                      |                      |                      | ADGM et al (2010)    |                      |                      |                      |
|-----------------|------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                 | Male                         |                      | Female               |                      | Male                 |                      | Female               |                      |
|                 | 1 <sup>st</sup>              | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      | 1 <sup>st</sup>      | 2 <sup>nd</sup>      |
| White           | 0.061***<br>(0.005)          |                      | 0.052***<br>(0.004)  |                      | 0.074***<br>(0.005)  | -                    | 0.052***<br>(0.004)  | -                    |
| Indian          | -0.095***<br>(0.009)         | 0.047***<br>(0.013)  | -0.126***<br>(0.008) | 0.026**<br>(0.011)   | -0.104***<br>(0.009) | 0.054***<br>(0.013)  | -0.137***<br>(0.008) | 0.026**<br>(0.011)   |
| Pakistani       | -0.275***<br>(0.014)         | -0.041**<br>(0.017)  | -0.153***<br>(0.020) | 0.010<br>(0.017)     | -0.269***<br>(0.013) | -0.028<br>(0.019)    | -0.147***<br>(0.020) | -0.009<br>(0.018)    |
| Black African   | -0.209***<br>(0.013)         | -0.128***<br>(0.025) | -0.171***<br>(0.011) | -0.073***<br>(0.024) | -0.243***<br>(0.014) | -0.173***<br>(0.028) | -0.208***<br>(0.011) | -0.085***<br>(0.025) |
| Black Caribbean | -0.203***<br>(0.014)         | -0.166***<br>(0.012) | -0.080***<br>(0.011) | -0.050***<br>(0.009) | -0.232***<br>(0.014) | -0.191***<br>(0.012) | -0.099***<br>(0.011) | -0.050***<br>(0.009) |
| Bangladeshi     | -0.567***<br>(0.023)         | -0.122**<br>(0.050)  | -0.153***<br>(0.036) | -0.025<br>(0.035)    | -0.574***<br>(0.022) | -0.146***<br>(0.051) | -0.185***<br>(0.039) | -0.046<br>(0.044)    |
| Chinese         | -0.086***<br>(0.026)         | 0.045<br>(0.033)     | -0.021<br>(0.020)    | 0.105***<br>(0.038)  | -0.058**<br>(0.026)  | 0.032<br>(0.040)     | -0.003<br>(0.020)    | 0.105**<br>(0.040)   |
| N               | 378,963                      | 358,023              | 400,028              | 378,153              | 331,043              | 317,275              | 347,006              | 332,569              |

Table reports coefficients obtained by regressing log hourly wages on generation and ethnicity dummies in an OLS; other covariates include a quartic in potential experience, region dummies and time dummies. \*\*\* indicates 1% level of significance; \*\* indicates 5% level of significance; and \* indicates 10% level of significance. Baseline: Native White Population. Robust standard errors as in ADGM. For the results in the text, standard errors are clustered by individual id.

**Table 4.16: Characteristics of the 2nd and 2.5th generations, mean values**

| Generation      | Age left full-time education |                   |                 |                   | Hourly earnings |                   |                 |                   | Employment probability |                   |                 |                   |
|-----------------|------------------------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-------------------|------------------------|-------------------|-----------------|-------------------|
|                 | <i>Males</i>                 |                   | <i>Females</i>  |                   | <i>Males</i>    |                   | <i>Females</i>  |                   | <i>Males</i>           |                   | <i>Females</i>  |                   |
|                 | 2 <sup>nd</sup>              | 2.5 <sup>th</sup> | 2 <sup>nd</sup> | 2.5 <sup>th</sup> | 2 <sup>nd</sup> | 2.5 <sup>th</sup> | 2 <sup>nd</sup> | 2.5 <sup>th</sup> | 2 <sup>nd</sup>        | 2.5 <sup>th</sup> | 2 <sup>nd</sup> | 2.5 <sup>th</sup> |
| White           | 18.0                         | 17.9              | 17.9            | 17.7              | 12.0            | 10.4              | 8.6             | 8.4               | 82.6                   | 80.4              | 69.8            | 68.1              |
| Indian          | 19.5                         | 16.3              | 19.6            | 16.7              | 9.6             | 6.1               | 7.6             | 5.6               | 69.6                   | 26.1              | 62.4            | 47.1              |
| Pakistani       | 19.3                         | 16.8              | 18.1            | 15.0              | 8.1             | 7.2               | 7.4             | 5.4               | 55.6                   | 45.5              | 34.9            | 36.8              |
| Black-African   | 21.8                         | 19.5              | 21.3            | 22.7              | 9.7             | 16.0              | 8.7             | 8.5               | 57.9                   | 69.2              | 65.3            | 37.5              |
| Black-Caribbean | 18.0                         | 17.8              | 18.4            | 18.2              | 9.7             | 9.5               | 9.5             | 7.6               | 71.6                   | 46.7              | 72.3            | 45.6              |
| Bangladeshi     | 18.0                         | 20.0              | 18.4            |                   | 7.0             | 20.9              | 8.4             |                   | 64.5                   | -                 | 38.9            | -                 |
| Chinese         | 20.8                         | -                 | 21.1            | 17.5              | 13.5            | 16.9              | 7.3             | 8.1               | 68.0                   | -                 | 62.5            | -                 |
| All             | 18.5                         | 17.9              | 18.4            | 17.7              | 10.8            | 10.5              | 8.5             | 8.3               | 73.9                   | 79.7              | 64.0            | 67.5              |

## Chapter 5: Conclusion

Economic theory suggests that emigration is driven by the incentive to increase one's utility. Referring to immigrants, Robert Winder (2013) note that "No one more gloriously passes the test proposed by Norman Tebbit's disdainful insistence that people should get on their bikes in search of work than those who have crossed the world in search of a better life." In the pursuit of increases in utility, immigrants end up contributing to the host country too (see Dustmann and Frattini (2014) who provide empirical evidence on the fiscal contribution of immigrants in the UK). Despite the various discourses on immigration in the period leading up to recent general elections in the UK, one common agreement has been that immigrants will continue to remain crucial for the British economy. This contribution largely rests on their integration in the labour market (and the general society) of the country.

Understanding how immigrants fare in the UK and in many other OECD has been the subject of increasing interest:

"Immigration and the integration of immigrants are also repeatedly mentioned as one of the main issues of concern in public opinion survey in many countries. At the same time, there are many preconceptions about the actual integration outcomes of immigrants and their children. Against this backdrop, having reliable facts is a prerequisite for a better informed public debate and for better targeted policy making." [OECD (2015), "Indicators of Immigrants Integration, Settling In."]

The present thesis contributes to the literature on the assimilation of immigrants in the UK via three self-contained essays, which examine their labour market outcomes and those of the children born to them in the country.

With the aim of controlling immigration, the UK has successively introduced various restrictive measures over the years. Today, while immigrants from certain countries require leave to enter and work in the UK, other do not. For example, nationals of member states of the EU can freely enter the UK and take up employment in any occupation just as the natives; American nationals and Nigerian nationals on the

other hand, for example, would require some form of work permit to enter the UK for employment purposes. While it is argued that the inferior labour market outcomes, including their occupational outcomes, of some groups of immigrants result from lower human capital levels, often associated with the place of origin of the immigrant, other have suggested that a host country may find it desirable that immigrants and natives have different occupational distributions.

The contribution of Chapter 2 has been to provide empirical evidence on the role of immigration policy on the occupational outcomes of immigrants. The exogenous change in the UK's immigration policy in May 2004 that allowed immigrants from A8 countries free access to the UK labour market overnight is used as a quasi-natural experiment to tease out the role of immigration policy on the occupational outcomes of immigrants from other factors that influence these outcome. Looking at a sample of recent immigrants, the chapter finds that the change in the UK's immigration policy led to a slight increase, although not statistically robust, in the odds of observing A8 immigrants who entered the UK after 2004 in elementary occupations relative to professional occupations. This finding is consistent with the theoretical prediction that reducing migration cost may not necessarily alter the skill composition of migrants, especially if this cost is constant or the extent of its heterogeneity across incoming migrants is relatively small.

To the best of our knowledge, this chapter is the first to examine the effect of a complete relaxation of immigration policy on the occupational outcomes of immigrants. The implementation of restrictive immigration policies are costly to the policy maker, employers as well as migrants. As the empirical evidence shows high rates of employment amongst A8 immigrants after 2004 and that they are more likely to be net contributors to the public purse (Dustamann and Frattini, 2014), our findings raise the question about how effective and important restrictive immigration policies are for the economy. In light of the prospect of new countries joining the EU, this study could be of interest to policy makers.

Chapter 2 has also explored the determinants of occupational outcomes of A8 immigrants, focusing on those who found their current employment after May 2004. The findings from this analysis suggest that duration of stay in the UK and income

levels at the source country of the immigrants play important roles in their occupational outcomes. The relative lower occupational attainment levels of some (highly educated) immigrants, in comparison to natives, despite lengthy time of stay in the host country, may possibly come from defining forgone income in term of the home country instead of the host country. If such is the cases, then there is room for Pareto improvement which would benefit both the immigrant and the host country; it would otherwise be a brain waste for both the host country and home country of the immigrant.

Chapter 3 has examined whether the occupational composition of a job seeker's network of social contacts has implications for their own occupational outcome. We have defined a network based on the country of birth of the immigrant and their region of work. The mechanism explored in the chapter is that referrals and job information, which are transmitted through network of social contacts are likely to be in the occupations of employed individuals within these networks. The chapter particularly explores whether these implications vary according to the length of stay of the immigrant in the UK, unlike in the scarce literature on the topic – i.e. Patel and Vella (2013). To the best of our knowledge, this chapter is the first to provide empirical evidence of the implications of the occupational composition of a network of social contacts on the occupational outcomes of the job seeker in the context of the UK.

A contribution of this chapter has been to show that the occupational composition of fellow co-ethnics has decreasing effect on the occupational outcomes of the job seeker with length of stay of the immigrant in the host country. This finding suggests that fellow co-ethnics may form the initial pool of contacts of a new immigrant, but that they potentially widen their circle of social contacts with time spent in the UK. It is possible that immigrants diversify their pool of social contacts as their stay in the UK lengthen, thereby undergoing a form of “social assimilation”. While we are unable to identify the country of birth of the person whom a job referral came from in the data used, it is reasonable to stipulate that the enlarged network of social contacts may include natives as well as immigrants from other countries different to

the job seeker's. This could also potentially shelter them from negative employment shocks amongst compatriots.

Chapter 4 has used a dataset with information on parental country of birth, unlike in previous studies, to obtain a more accurate measure of second-generation migrant status and examine relative labour market outcomes of second generation immigrants across all major ethnic groups in the UK. The chapter shows that it is important to correctly distinguish second generation immigrants from natives when examining any intergenerational mobility in the labour market outcomes of immigrants. Because of data limitations, previous studies on the subject have commonly relied on ethnicity and birth in the UK as a proxy measure for the nativity status of individuals. As such, all UK-born of non-White ethnicity are classified as second generation immigrants regardless of the place of birth of their parents. This method leads to errors of classification in regressors, which we show that, even at trivial level, can generate large levels of uncertainty in the estimated coefficients of interest. This is particularly the case for ethnic minority groups with a long history of residence in the UK.

In accordance with other studies, the chapter finds that most groups of immigrants and the children born to them in the UK generally have lower employment probabilities and earnings than their native counterparts, despite their educational attainments advantage. There is, nevertheless, a clear improvement in these outcomes across generations amongst most ethnic groups. The presence of potential classification errors in the data, however, can lead to a less clear picture of intergenerational mobility in relative labour market outcomes when using LFS data: while the children of different groups of immigrants might have better labour market than their respective parents, relative to natives, it is also quite possible they might have worse labour market outcomes too.

A question of interest would be to examine how the labour market outcomes of UK-born individuals with a foreign-born parent and a UK-born parent compare to those of individuals with both UK-born parents. Chapter 4 does not deal with this question, yet the answer to this would also add to our understanding of the integration of immigrants and the children born to them in the country. Closely

related to this is whether the labour market outcomes of immigrants who entered the UK as children differ to those of natives across different ethnic groups given individuals who enter the UK as children are less likely to lack host country-specific human capital.

Further research related to Chapter 2 could examine how the change in policy affected the occupational outcomes of different cohorts of A8 immigrants who entered the UK after 2004, and their assimilation process too. We have only looked at its effects on recent immigrants who entered the UK shortly after 2004. Also, it would be interesting to look at how the annulment of immigration restrictions faced by A8 nationals affected vacancy rates in different occupations in subsequent years to the year the change in immigration policy occurred. This may shed light on the role immigration policy aimed at compensating for shortages in the labour market of the host country.

Very little evidence currently exists on the effect of institutional constraints on the likelihood of finding employment through informal channels such as referrals, despite their instrumental role in matching job seekers and employers. For example, an institutional constraint such as immigration restrictions could, at the margin, arguably lead to employers having preference for immigrants who face no such constraints. If this is the case, freedom of movement of labour agreements such as the one amongst member states of the EU may alleviate lengths of unemployment spells.

## Bibliography

- Aigner, D. J. (1973). Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics*, 1(1):49–59.
- Alesina, A., Harnoss, J., and Rapoport, H. (2013). Birthplace diversity and economic prosperity. *NBER Working Paper*.
- Alesina, A., Harnoss, J., and Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2):101–138.
- Algan, Y., Dustmann, C., Glitz, A., and Manning, A. (2010). The economic situation of first and second-generation immigrants in France, Germany and the United Kingdom. *The Economic Journal*, 120(542):F4–F30.
- Amuedo-Dorantes, C. and De la Rica, S. (2007). Labour market assimilation of recent immigrants in Spain. *British Journal of Industrial Relations*, 45(2):257–284.
- Andersson, F., Burgess, S., and Lane, J. (2009). Do as the neighbors do: The impact of social networks on immigrant employment. *CMPO Working Paper*.
- Angrist, J. D., and Pischke, J. S. (2009). Mostly Harmless Econometrics: an Empiricist’s Companion. *Princeton: Princeton University Press*.
- Antecol, H., Cobb-Clark, D. A., and Trejo, S. J. (2003). Immigration policy and the skills of immigrants to Australia, Canada, and the United States. *Journal of Human Resources*, 38(1):192–218.
- Aydemir, A. (2011). Immigrant selection and short-term labor market outcomes by visa category. *Journal of Population Economics*, 24(2):451–475.
- Aydemir, A., Chen, W. H., and Corak, M. (2009). Intergenerational earnings mobility among the children of Canadian immigrants. *The Review of Economics and Statistics*, 91(2):377–397.
- Aydemir, A. and Sweetman, A. (2007). Educational attainment and labour market outcomes of the first and children of migrants: A comparison of the US and Canada. *Research in Labor Economics*, 27:215–270.
- Barrett, A. and Duffy, D. (2008). Are Ireland’s immigrants integrating into its labor market? *International Migration Review*, 42(3):597–619.
- Battu, H., Seaman, P., and Zenou, Y. (2011). Job contact networks and the ethnic minorities. *Labour Economics*, 18(1):48–56.

- Bayer, P., Ross, S. L., and Topa, G. (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116(6):1150–1196.
- Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US. *The Review of Economic Studies*, 79(1):128–161.
- Beine, M., Docquier, F., and Ozden, C. (2011). Diasporas. *Journal of Development Economics*, 95(1):30–41.
- Bell, B., Blundell, R., and Van Reenen, J. (1999). Getting the unemployed back to work: the role of targeted wage subsidies. *International Tax and Public Finance*, 6(3):339–360.
- Bell, B. D. (1997). The performance of immigrants in the United Kingdom: evidence from the GHS. *The Economic Journal*, 107(441):333–344.
- Belot, M. V. and Hatton, T. J. (2012). Immigrant selection in the OECD. *The Scandinavian Journal of Economics*, 114(4):1105–1128.
- Bentolila, S., Michelacci, C., and Suarez, J. (2010). Social contacts and occupational choice. *Economica*, 77(305):20–45.
- Bertoli, S. and Rapoport, H. (2015). Heaven’s swing door: Endogenous skills, migration networks, and the effectiveness of quality-selective immigration policies. *The Scandinavian Journal of Economics*, 117(2):565–591.
- Blanchflower, D. G. and Shadforth, C. (2009). Fear, unemployment and migration. *The Economic Journal*, 119(535):F136–F182.
- Blundell, R., Dias, M. C., Meghir, C., and van Reenen, J. (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association*, 2(4):569–606.
- Bollinger, C. R. (1996). Bounding mean regressions when a binary regressor is mismeasured. *Journal of Econometrics*, 73(2):387–399.
- Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, 3(4):463–489.
- Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, 77(4):531–553.

- Borjas, G. J. (1989). Economic theory and international migration. *International Migration Review*, 23(3):457–485.
- Borjas, G. J. (1993). Immigration policy, national origin, and immigrant skills: A comparison of Canada and the United States. In *Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States*, pages 21–44. University of Chicago Press.
- Borjas, G. J. (1994a). The economics of immigration. *Journal of Economic Literature*, 32(4):1667–1717.
- Borjas, G. J. (1994b). Long-run convergence of ethnic skill differentials: The children and grandchildren of the great migration. *Industrial & Labor Relations Review*, 47(4):553–573.
- Borjas, G. J. (1999). The economic analysis of immigration. *Handbook of Labor Economics*, 3:1697–1760.
- Boskin, M. J. (1974). A conditional logit model of occupational choice. *Journal of Political Economy*, 82(2):389–398.
- Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. *Handbook of Econometrics*, 5:3705–3843.
- Bramoullé, Y. and Saint-Paul, G. (2010). Social networks and labor market transitions. *Labour Economics*, 17(1):188–195.
- Cahuc, P. and Fontaine, F. (2009). On the efficiency of job search with social networks. *Journal of Public Economic Theory*, 11(3):411–439.
- Calvo-Armengol, A. and Jackson, M. O. (2004). The effects of social networks on employment and inequality. *The American Economic Review*, 94(3):426–454.
- Calvo'-Armengol, A. and Zenou, Y. (2005). Job matching, social network and word-of-mouth communication. *Journal of Urban Economics*, 57(3):500–522.
- Cameron, A. and Trivedi, P. (2009). Microeconometrics using stata. *Volume 5, Stata press College Station, TX*.
- Campbell, S. (2013). Over-education among A8 migrants in the UK. *Institute of Education, DoQSS Working Paper*, (13-09).

- Cappellari, L. and Tatsiramos, K. (2010). With a little help from my friends? The effects of friendship ties on unemployment transitions. *University of Leicester Working Paper*.
- Cappellari, L. and Tatsiramos, K. (2015). With a little help from my friends? Quality of social networks, job finding and job match quality. *European Economic Review*, 78:55–75.
- Chiquiar, D. and Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy*, 113(2):239–281.
- Chiswick, B. R. (1978). The effect of Americanisation on the earnings of foreign-born men. *The Journal of Political Economy*, 86(5):897–921.
- Chiswick, B. R. (1980). The earnings of white and coloured male immigrants in Britain. *Economica*, 47(185):81–87.
- Chiswick, B. R., Lee, Y. L., and Miller, P. W. (2005). A longitudinal analysis of immigrant occupational mobility: A test of the immigrant assimilation hypothesis1. *International Migration Review*, 39(2):332–353.
- Chuhay, R. (2013). Labor market and search through personal contacts. *The BE Journal of Theoretical Economics*, 13(1):191–213.
- Clark, K. and Drinkwater, S. (2008). The labour-market performance of recent migrants. *Oxford Review of Economic Policy*, 24(3):495–516.
- Czaika, M. and Parsons, C. (2015). The gravity of high-skilled migration policies. *IMI Policy Paper*.
- Drinkwater, S., Eade, J., and Garapich, M. (2009). Poles apart? EU enlargement and the labour market outcomes of immigrants in the United Kingdom. *International Migration*, 47(1):161–190.
- Dustmann, C. and Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, 113(489):695–717.
- Dustmann, C. and Faber, F. (2005). Immigrants in the British labour market. *Fiscal Studies*, 26(4):423–470.
- Dustmann, C. and Frattini, T. (2014). The fiscal effects of immigration to the UK. *The Economic Journal*, 124(580):F593–F643.

- Dustmann, C., Glitz, A., and Schoenberg, U (2011). Referral-based job search networks. *Norface Research Programme on Migration, Department of Economics, University College London*, No. 2011012.
- Dustmann, C. and Kirchkamp, O. (2002). The optimal migration duration and activity choice after re-migration. *Journal of Development Economics*, 67(2):351-372.
- Dustmann, C. and Theodoropoulos, N. (2010). Ethnic minority immigrants and their children in Britain. *Oxford Economic Papers*, 62(2):209–233.
- Dustmann, C. and Weiss, Y. (2007). Return migration: theory and empirical evidence from the UK. *British Journal of Industrial Relations*, 45(2):236–256.
- Engels, F. (1845). The condition of the working class in England. *The Institute of Marxism-Leninism*
- Fontaine, F. (2007). A simple matching model with social networks. *Economics Letters*, 94(3):396–401.
- Fontaine, F. (2008). Why are similar workers paid differently? The role of social networks. *Journal of Economic Dynamics and Control*, 32(12):3960–3977.
- Frijters, P., Shields, M. A., and Price, S. W. (2005). Job search methods and their success: a comparison of immigrants and natives in the UK. *The Economic Journal*, 115(507):F359–F376.
- Gagliardi, L. and Lemos, S. (2016). Evidence on immigrants’ assimilation into recipient labour markets using UK longitudinal data between 1981 and 2006. *Journal of Economic Geography*, 16(3):547–583.
- Galenianos, M. (2014). Hiring through referrals. *Journal of Economic Theory*, 152:304–323.
- Garber, S. and Klepper, S. (1980). Extending the classical normal errors-in-variables model. *Econometrica*, 48(6):1541-1546 1541–1546.
- Giulietti, C., Schluter, C., and Wahba, J. (2013). With a lot of help from my friends: Social networks and immigrants in the UK. *Population, Space and Place*, 19(6):657–670.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360-1380.

- Green, A. G. and Green, D. A. (1995). Canadian immigration policy: The effectiveness of the point system and other instruments. *Canadian Journal of Economics*, 28(4b):1006–1041.
- Green, D. A. (1999). Immigrant occupational attainment: Assimilation and mobility over time. *Journal of Labor Economics*, 17(1):49–79.
- Gundersen, C. and Kreider, B. (2009). Bounding the effects of food insecurity on children’s health outcomes. *Journal of Health Economics*, 28(5):971–983.
- Heath, A. and Li, Y. (2008). Period, life-cycle and generational effects on ethnic minority success in the British labour market. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 48:277–306.
- Heath, A. F., Rethon, C., and Kilpi, E. (2008). The second generation in Western Europe: Education, unemployment, and occupational attainment. *Annual Review of Sociology*, 34:211–235.
- Holzer, H. J. (1988). Search method use by unemployed youth. *Journal of Labor Economics*, 6(1):1–20.
- Horvath, G. (2014). Occupational mismatch and social networks. *Journal of Economic Behavior & Organization*, 106:442–468.
- Ioannides, Y. M. and Datcher Loury, L. (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 42(4):1056–1093.
- Kanas, A., Chiswick, B. R., Lippe, T., and Tubergen, F. (2012). Social contacts and the economic performance of immigrants: a panel study of immigrants in Germany. *International Migration Review*, 46(3):680–709.
- Klepper, S. (1988). Regressor diagnostics for the classical errors-in-variables model. *Journal of Econometrics*, 37(2):225–250.
- Kreider, B. (2010). Regression coefficient identification decay in the presence of infrequent classification errors. *The Review of Economics and Statistics*, 92(4):1017–1023.
- Kugler, A., Jimeno, J. F., and Hernanz, V. (2002). Employment consequences of restrictive permanent contracts: Evidence from Spanish labor market reforms. *CEPR Discussion Papers*
- Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends (R) in Econometrics*, 4(3):165–224.

- Lee, M. J. and Kang, C. (2006). Identification for difference in differences with cross-section and panel data. *Economics letters*, 92(2):270–276.
- Li, Y. (2010). The labour market situation of minority ethnic groups in Britain and the US. *EurAmerica: A Journal of European and American Studies*, 40(2):259–309.
- Longhi, S. and Rokicka, M. (2012). European immigrants in the UK before and after the 2004 enlargement: Is there a change in immigrant self-selection? *Technical report, Norface Research Programme on Migration, Department of Economics, University College London*.
- Mahuteau, S. and Junankar, P. R. (2008). Do migrants get good jobs in Australia? The role of ethnic networks in job search. *Economic Record*, 84(s1):S115–S130.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.
- Mattoo, A., Neagu, I. C., and Ozden, C. (2008). Brain waste? Educated immigrants in the US labor market. *Journal of Development Economics*, 87(2):255–269.
- Macdonald, I. A. and Webber, F. (2005). Macdonald's Immigration Law and Practice. *Butterworths Law*, 6<sup>th</sup> Edition
- McFadden, Daniel. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics*, ed. by P. Zarembka, New York: Academic Press, 1974, 105-142.
- McKenzie, D. and Rapoport, H. (2010). Self-selection patterns in Mexico-US migration: the role of migration networks. *The Review of Economics and Statistics*, 92(4):811–821.
- Mesnard, A. (2004). Temporary migration and capital market imperfections. *Oxford Economic Papers*, 56(2):242–262.
- Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13(2):151–161.
- Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal migration in the United States. *The Journal of Economic Perspectives*, 25(3):173–196.
- Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American economic review*, 81(5):1408-1418.
- Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, 118(2):549-599.

- OECD and EU. (2015). Indicators of Immigrant Integration 2015: Settling In. *OECD Publishing, Paris*.
- Patacchini, E. and Zenou, Y. (2012). Ethnic networks and employment outcomes. *Regional Science and Urban Economics*, 42(6):938–949.
- Patel, K. and Vella, F. (2007). Immigrant networks and their implications for occupational choice and wages. *IZA Discussion Papers No 3217*.
- Patel, K. and Vella, F. (2013). Immigrant networks and their implications for occupational choice and wages. *Review of Economics and Statistics*, 95(4):1249–1277.
- Platt, L. (2005). The intergenerational social mobility of minority ethnic groups. *Sociology*, 39(3):445–461.
- Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear difference-in-differences models. *Economics Letters*, 115(1):85–87.
- Ramakrishnan, K. (2004). Second-Generation Immigrants? The “2.5 Generation” in the United States. *Social Science Quarterly*, 85(2):380–399
- Rees, A. (1966). Information networks in labor markets. *The American Economic Review*, 56(1/2):559–566.
- Rooth, D. O. and Saarela, J. (2007). Selection in migration and return migration: Evidence from micro data. *Economics Letters*, 94(1):90–95.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.
- Saleheen, J. and Shadforth, C. (2006). The economic characteristics of immigrants and their impact on supply. *Bank of England Quarterly Bulletin, Winter*.
- Salt, J., Okolski, M., et al. (2014). Polish emigration to the UK after 2004; why did so many come? *Central and Eastern European Migration Review*, (02):11–37.
- Spencer, S. (2011). The migration debate. *Policy Press*.
- Supperstone, M., Laing, E., O'dempsey, D. & O'dempsey, D. (1994). Immigration: The Law and Practice. *Longman Practitioner*, 3<sup>rd</sup> Edition.
- Stupnytska, Y. and Zaharieva, A. (2015). Explaining u-shape of the referral hiring pattern in a search model with heterogeneous workers. *Journal of Economic Behavior & Organization*, 119:211–233.

Sumption, M. (2009). Social networks and polish immigration to the UK. *Economics of Migration Working Paper*.

Wahba, J. (2015). Selection, selection, selection: the impact of return migration. *Journal of Population Economics*, 28(3):535–563.

Wahba, J. and Zenou, Y. (2005). Density, social networks and job search methods: Theory and application to Egypt. *Journal of Development Economics*, 78(2):443–473.

Winder, R. (2013). Bloody foreigners: The story of immigration to Britain. *Hachette, UK*.

Zaharieva, A. (2016). On the puzzle of diversification in social networks with occupational mismatch. *Centre for Mathematical Economics (IMW) Working Paper*.